



Year: 2021

Autosomal recessive loci contribute significantly to quantitative variation of male fertility in a dairy cattle population

Hiltpold, Maya ; Kadri, Naveen Kumar ; Janett, Fredi ; Witschi, Ulrich ; Schmitz-Hsu, Fritz ; Pausch, Hubert

Abstract: Background: Cattle are ideally suited to investigate the genetics of male fertility. Semen from individual bulls is used for thousands of artificial inseminations for which the fertilization success is monitored. Results from the breeding soundness examination and repeated observations of semen quality complement the fertility evaluation for each bull. Resultat: In a cohort of 3881 Brown Swiss bulls that had genotypes at 683,609 SNPs, we reveal four novel recessive QTL for male fertility on BTA1, 18, 25, and 26 using haplotype-based association testing. A QTL for bull fertility on BTA1 is also associated with sperm head shape anomalies. All other QTL are not associated with any of the semen quality traits investigated. We perform complementary fine-mapping approaches using publicly available transcriptomes as well as whole-genome sequencing data of 125 Brown Swiss bulls to reveal candidate causal variants. We show that missense or nonsense variants in SPATA16, VWA3A, ENSBTAG00000006717 and ENSBTAG00000019919 are in linkage disequilibrium with the QTL. Using whole-genome sequence data, we detect strong association ($P = 4.83 \times 10^{12}$) of a missense variant (p.Ile193Met) in SPATA16 with male fertility. However, non-coding variants exhibit stronger association at all QTL suggesting that variants in regulatory regions contribute to variation in bull fertility. Conclusion: Our findings in a dairy cattle population provide evidence that recessive variants may contribute substantially to quantitative variation in male fertility in mammals. Detecting causal variants that underpin variation in male fertility remains difficult because the most strongly associated variants reside in poorly annotated non-coding regions.

DOI: <https://doi.org/10.1186/s12864-021-07523-3>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-210727>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Hiltpold, Maya; Kadri, Naveen Kumar; Janett, Fredi; Witschi, Ulrich; Schmitz-Hsu, Fritz; Pausch, Hubert (2021). Autosomal recessive loci contribute significantly to quantitative variation of male fertility in a dairy cattle population. BMC Genomics, 22(1):225.

DOI: <https://doi.org/10.1186/s12864-021-07523-3>

RESEARCH ARTICLE

Open Access



Autosomal recessive loci contribute significantly to quantitative variation of male fertility in a dairy cattle population

Maya Hiltbold^{1*}, Naveen Kumar Kadri¹, Fredi Janett², Ulrich Witschi³, Fritz Schmitz-Hsu³ and Hubert Pausch¹

Abstract

Background: Cattle are ideally suited to investigate the genetics of male fertility. Semen from individual bulls is used for thousands of artificial inseminations for which the fertilization success is monitored. Results from the breeding soundness examination and repeated observations of semen quality complement the fertility evaluation for each bull.

Results: In a cohort of 3881 Brown Swiss bulls that had genotypes at 683,609 SNPs, we reveal four novel recessive QTL for male fertility on BTA1, 18, 25, and 26 using haplotype-based association testing. A QTL for bull fertility on BTA1 is also associated with sperm head shape anomalies. All other QTL are not associated with any of the semen quality traits investigated. We perform complementary fine-mapping approaches using publicly available transcriptomes as well as whole-genome sequencing data of 125 Brown Swiss bulls to reveal candidate causal variants. We show that missense or nonsense variants in *SPATA16*, *VWA3A*, *ENSBTAG00000006717* and *ENSBTAG00000019919* are in linkage disequilibrium with the QTL. Using whole-genome sequence data, we detect strong association ($P = 4.83 \times 10^{-12}$) of a missense variant (p.Ile193Met) in *SPATA16* with male fertility. However, non-coding variants exhibit stronger association at all QTL suggesting that variants in regulatory regions contribute to variation in bull fertility.

Conclusion: Our findings in a dairy cattle population provide evidence that recessive variants may contribute substantially to quantitative variation in male fertility in mammals. Detecting causal variants that underpin variation in male fertility remains difficult because the most strongly associated variants reside in poorly annotated non-coding regions.

Keywords: Bull fertility, Quantitative trait loci, Semen quality, Association study, Haplotype

Background

Male fertility is a complex trait that is determined by genetic and non-genetic sources of variation. Because of its low heritability, a large sample size is required to investigate the genetic architecture of male fertility. Large cohorts of males with repeated measurements for reproductive traits are not available in many species including

humans, but are accessible in cattle where semen samples from individual bulls are used for thousands of artificial inseminations [1]. Because these data facilitate the disentanglement of male and female factors contributing to establishing pregnancy [2], the fertility of bulls can be quantified objectively.

Semen quality has a large effect on insemination success [3–7]. On entering the semen collection center, the semen quality of each bull is examined as part of an in-depth breeding soundness evaluation [8]. The semen quality of artificial insemination bulls varies over time,

* Correspondence: maya.hiltbold@usys.ethz.ch

¹Animal Genomics, Institute of Agricultural Sciences, ETH Zürich, Eschikon 27, 8315 Lindau, Switzerland

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

consequently the fertilization rates may differ substantially between ejaculates. In order to ensure high and uniform insemination success, parameters such as ejaculate volume, sperm concentration, sperm head and tail morphology, and sperm motility are examined in all ejaculates immediately after semen collection. Ejaculates that fulfill predefined quality requirements are diluted and filled into doses that contain between 15 and 25 million sperm, depending on semen quality. After subsequent cryopreservation, at least the frozen-thawed sperm motility is checked. Computer-assisted sperm analysis or flow cytometric assays can add additional information on the semen quality [9]. The combination of multiple semen quality parameters may enable predicting bull fertility to a certain extent [10–14].

Phenotypes describing male reproductive traits and dense microarray-derived genotypes for many artificial insemination sires enable genome-wide association testing. Quantitative trait loci (QTL) for bull fertility have been detected in Holstein [15–19] and Jersey [20] cattle. Nevertheless, these studies did not attempt to reveal the causal variants underlying the QTL with whole-genome sequence information. Moreover, except in [19], the sample sizes were relatively small providing low statistical power to detect QTL for a trait with low to moderate heritability like bull fertility. Recent investigations provide growing evidence that non-additive effects contribute to inherited variation in bull fertility [16, 21–24]. However, the underlying genetic mechanisms remain largely unknown.

Here, we apply haplotype-based association testing to detect QTL for bull fertility and semen quality in a large mapping cohort of bulls from the Brown Swiss cattle breed that had dense microarray-derived genotypes. Our association studies revealed five recessive QTL for bull fertility. Using two complementary sequence-based fine-mapping approaches and transcriptome data, we identify candidate causal variants in coding sequences of genes that are expressed in the male reproductive tract.

Results

To detect QTL for bull fertility, we considered 3736 Brown Swiss (BSW) bulls that were kept at semen collection centers in Switzerland, Germany and Austria. Bull fertility was quantified based on artificial insemination success adjusted for confounding factors (see Methods). The bulls had partially imputed genotypes at 589,791 autosomal SNPs with minor allele frequency greater than 0.5%. Genomic restricted maximum likelihood estimation indicated that the autosomal markers explained $0.10 \pm 0.02\%$ of the phenotypic variance in bull fertility.

Recessive QTL for bull fertility are located on chromosomes 1, 6, 18, 25 and 26

Haplotype-based association tests that were based on an additive mode of inheritance revealed three QTL located on chromosomes 1, 6, and 26 for bull fertility that exceeded the Bonferroni-corrected genome-wide significance threshold of 2.68×10^{-7} (Table 1, Fig. 1a).

We previously showed that the QTL for bull fertility at BTA6 is in linkage disequilibrium with the recessive BTA6:58,373,887 T-allele (rs474302732) that activates cryptic splicing in *WDR19*, resulting in reduced semen quality [22]. In order to investigate the association of the BTA6:58,373,887 T-allele in our mapping cohort, we imputed the genotypes of the 3736 bulls to the whole-genome sequence level using a breed-specific reference panel of 368 sequenced cattle. A sequence-based association analysis that was based on a recessive mode of inheritance confirmed that rs474302732 is strongly associated ($P = 2.72 \times 10^{-45}$) with bull fertility. The association with bull fertility was stronger for rs474302732 than the most significantly associated haplotype (Table 1). Four non-coding variants (57,900,948 bp, 57,950,075 bp, 57,408,389 bp, 57,810,181 bp) had slightly lower P values ($P = 5.36 \times 10^{-46}$ - 3.81×10^{-47}) than rs474302732. Due to its large effect on male fertility, we subsequently fixed the top haplotype of the BTA6 QTL as a covariate in the association model.

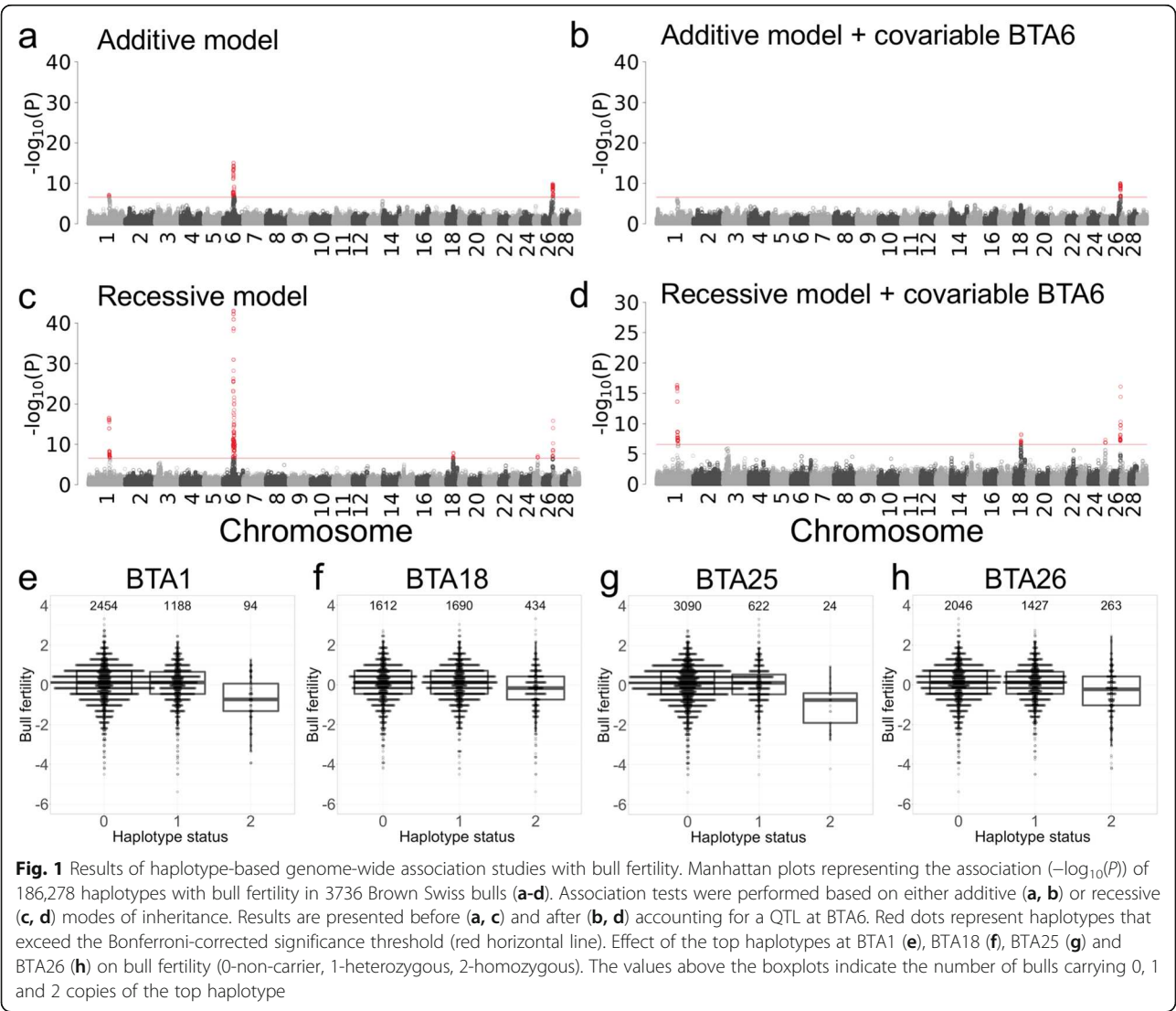
When the top haplotype of the BTA6 QTL was fixed as a covariate in the model, the association signal at BTA6 disappeared (Fig. 1b). In the conditional analysis, the P value of the top-associated haplotype at the BTA1 QTL did not meet the Bonferroni-corrected significance threshold by a small margin (Fig. 1b). The QTL on BTA26 remained significant ($P = 9.96 \times 10^{-11}$).

Next, we repeated the haplotype-based association test assuming a recessive mode of inheritance. The association with bull fertility was more pronounced (i.e., the P values of the top haplotypes were lower) at the QTL on chromosomes 1, 6 and 26 (Table 1, Fig. 1c & d). Moreover, the recessive model revealed two QTL at chromosomes 18 and 25 that were not detected using the additive model. Visual inspection of the haplotype effects corroborated that the four novel QTL detected on chromosomes 1, 18, 25, and 26 act in a recessive manner (Fig. 1e, f, g, h). Compared to non-carrier and heterozygous haplotype carrier bulls, the fertility of bulls that carry the BTA1, BTA18, BTA25 and BTA26 top haplotype in the homozygous state is reduced by 0.80 ± 0.10 , 0.28 ± 0.05 , 1.09 ± 0.20 and 0.52 ± 0.06 phenotypic standard deviations, respectively. A linear regression analysis conditional on the top 10 principal components and the top haplotype at the BTA6 QTL confirmed that fertility does not differ between heterozygous and non-carrier bulls at the four novel QTL ($P_{\text{BTA1}} = 0.29$, $P_{\text{BTA18}} = 0.61$,

Table 1 Position of five recessive QTL for male fertility in BSW cattle

Additive mode of inheritance					Recessive mode of inheritance			
	Start position	Stop position	Frequency of the fertility-decreasing haplotype	P value	Start position	Stop position	Frequency of the fertility-decreasing haplotype	P value
BTA1 ^a				ns	93,830,840	94,035,932	0.18	4.16 × 10 ⁻¹⁷
BTA6	57,949,492	58,100,971	0.26	8.51 × 10 ⁻¹⁶	57,949,492	58,100,971	0.26	9.34 × 10 ⁻⁴⁴
BTA18 ^a				ns	36,256,440	36,555,965	0.34	5.38 × 10 ⁻⁹
BTA25 ^a				ns	19,628,841	19,834,098	0.09	4.56 × 10 ⁻⁸
BTA26 ^a	49,567,535	49,720,699	0.08	9.96 × 10 ⁻¹¹	50,746,717	50,993,657	0.26	7.53 × 10 ⁻¹⁷

^aresults are based on a conditional association analysis where the BTA6 top haplotype was fixed as covariate in the model



$P_{\text{BTA25}} = 0.79$, $P_{\text{BTA26}} = 0.59$), confirming their recessive inheritance.

Between 24 and 434 out of the 3736 bulls were homozygous carriers of fertility-decreasing haplotypes at the four novel and one previously detected QTL. Across the five QTL, 2827, 604, 259, 44, and 2 bulls were homozygous carriers of 0, 1, 2, 3, and 4 top haplotypes, respectively. In order to quantify the impact of the five recessive QTL on insemination success in the BSW population, we investigated the 56-day non-return rate (NRR56; mean = $65.06 \pm 4.53\%$) in cows inseminated with semen from 1322 BSW bulls, estimated from at least 300 inseminations per bull. The NRR56 is the proportion of cows that is not re-inseminated within a 56-day interval after the first insemination and is considered as a proxy for the insemination success. The mean NRR56 is 0.5 phenotypic standard deviations higher ($P = 2.75 \times 10^{-8}$, two-tailed t-test) in 1000 bulls that do not carry any of the top haplotypes in the homozygous state than in 322 bulls that are homozygous for at least one of the five recessive QTL (65.48 ± 4.33 vs. $63.76 \pm 4.89\%$).

A bull fertility QTL on BTA1 affects sperm head morphology

A recessive QTL for bull fertility is located on BTA1. The top haplotype ($P = 4.16 \times 10^{-17}$) resides between 93,830,840 and 94,035,932 bp. The fertility-decreasing haplotype occurs at a frequency of 0.18 in the BSW population. We detected 94 and 1188 bulls that carried the top haplotype in the homozygous and heterozygous state, respectively.

In order to unravel the mechanism through which the QTL impacts male fertility, we analysed routinely collected semen quality data that were available for 32 homozygous, 372 heterozygous and 498 non-carrier bulls. The BTA1 QTL was not associated with ejaculate volume ($P = 0.07$), sperm concentration ($P = 0.19$), sperm motility ($P = 0.10$), sperm head morphology ($P = 0.53$) or sperm tail morphology ($P = 0.19$, Supplementary Figure 1) assessed in fresh semen. However, the insemination straws contained more ($+ 1.57$ million, $P = 4.78 \times 10^{-6}$, Supplementary Figure 1f) sperm in homozygous than heterozygous and non-carrier bulls, which indicates an attempt to compensate for low sperm quality by providing additional sperm per straw, although the association with the number of sperm per straw did not meet the Bonferroni-corrected significance threshold ($P_{\text{Bonf}} = 2.68 \times 10^{-7}$) by a small margin. Moreover, more ejaculates were discarded due to insufficient semen quality from homozygous than heterozygous and non-carrier bulls (14.8 vs. 6.5%).

The association pattern of the BTA1 top haplotype was puzzling. Homozygosity for the top haplotype

compromises bull fertility but not routinely recorded fresh semen quality. However, the insemination doses of homozygous bulls contain more sperm per straw. In order to further examine this apparent contradiction, we studied 1577 sperm morphology evaluations that were collected as part of the andrological examination of 575 prospective artificial insemination bulls. During the andrological examination, sperm morphology of at least one ejaculate per bull is systematically evaluated, before the semen is used for artificial inseminations. This evaluation is repeated either until the semen quality meets predefined quality thresholds or the bull is rejected for artificial insemination due to insufficient semen quality. Sperm morphology evaluations were provided for between 1 and 13 (median = 2) ejaculates of 575 Swiss BSW bulls. From these reports, we derived 21 sperm quality phenotypes (see Methods).

Genome-wide haplotype-based association testing revealed that the QTL at BTA1 affects four (out of 21) sperm morphology features. Bulls that are homozygous for the top haplotype produce ejaculates that contain less normal spermatozoa ($- 10.6\%$, $P = 1.72 \times 10^{-10}$), twice as many sperm with non-compensatory defects ($+ 5.4\%$, $P = 9.07 \times 10^{-16}$), 1.5 times more sperm with major defects ($+ 8.9\%$, $P = 2.55 \times 10^{-11}$) and twice as many sperm with head shape anomalies ($+ 5.6\%$, $P = 3.98 \times 10^{-17}$) (Fig. 2, Supplementary Figure 1 h, i, l, & m). These data suggest that an increased proportion of sperm with abnormal head morphology compromises bull fertility, as no other sperm morphology feature was affected by the BTA1 QTL. The number of sperm morphology evaluations was also higher ($P = 1.63 \times 10^{-10}$) for homozygous than heterozygous and non-carrier bulls (Supplementary Figure 1 g). Eventually, only 21 out of 34 homozygous bulls (61.7%) compared to 520 out of 561 (92.7%) heterozygous and non-carrier bulls passed the sperm morphology examination and produced ejaculates that are suitable for breeding ($\geq 65\%$ normal sperm and $\leq 20\%$ non-compensatory defects). Sperm morphology did not differ between heterozygous haplotype carriers and non-carrier bulls, corroborating recessive inheritance.

We examined and imaged fresh sperm from two bulls homozygous for the top haplotype using oil-immersion phase-contrast light microscopy. At semen collection, the bulls were 644 and 504 days old. Some spermatozoa had morphological anomalies (i.e., rounder and shorter heads, heads with rounder frontal part (pear-shape/pyriform), narrower head base (tapered), and heads with abnormal contour (uneven shaped)) that are classified as head shape anomalies in the sperm morphology evaluation (Supplementary Figures 2 & 3). However, the vast majority of the spermatozoa (85 and 84.6%) were normal. Acrosome defects were not apparent.

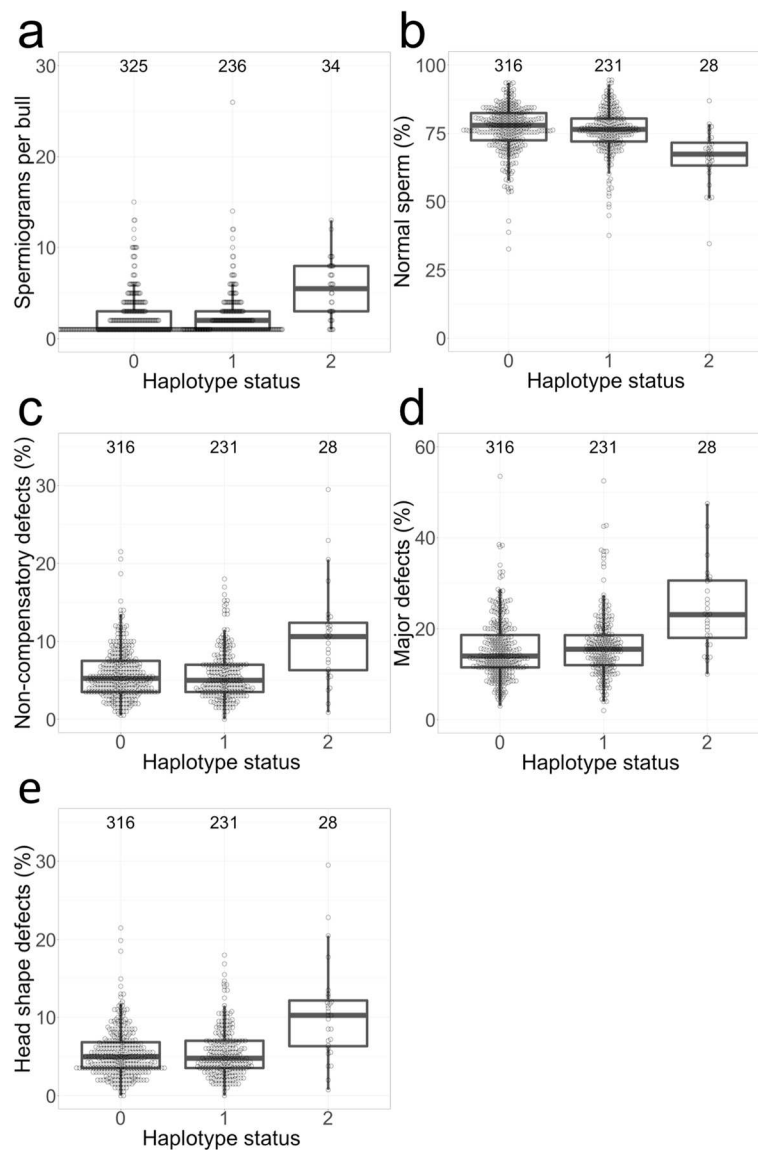


Fig. 2 Effect of the BTA1 top haplotype on sperm morphology. Boxplots representing the sperm morphology of non-carrier, heterozygous, and homozygous (haplotype status 0, 1, and 2) bulls. The numbers above the boxplots indicate the number of bulls in the respective group. Compared to non-carrier and heterozygous bulls, **a** the number of sperm morphology examinations is twice as high (5.32 ± 3.25 vs. 2.49 ± 2.50), **b** the amount of normal spermatozoa is ~ 10 percentage points lower ($66.19 \pm 10.43\%$ vs. $76.45 \pm 8.51\%$), **c** the amount of sperm with non-compensatory defects is ~ 5 percentage points higher ($10.61 \pm 6.28\%$ vs. $5.79 \pm 3.19\%$), **d** the amount of sperm with major defects is ~ 8 percentage points higher ($24.08 \pm 8.99\%$ vs. $15.70 \pm 6.82\%$), **e** and the amount of spermatozoa with head shape anomalies is ~ 5 percentage points higher ($10.49 \pm 6.30\%$ vs. $5.42 \pm 3.09\%$) in homozygous bulls

A missense variant in and non-coding variants upstream *SPATA16* are associated with the BTA1 QTL

In order to detect candidate causal variants for the increased number of abnormal sperm and reduced fertility associated with the BTA1 QTL, we applied a two-step fine-mapping approach. First, we used whole-genome sequence data of 125 BSW bulls to identify variants that are compatible with the inheritance of the top haplotype. The average fold coverage of the sequenced bulls was 10.1 ± 4.3 -fold. The BTA1 QTL top haplotype was

determined for the sequenced bulls based on their microarray-derived genotypes. Second, we imputed whole-genome sequence variant genotypes for the 3736 BSW bulls of the mapping cohort using 368 reference animals to perform an association study between imputed sequence variant genotypes and bull fertility.

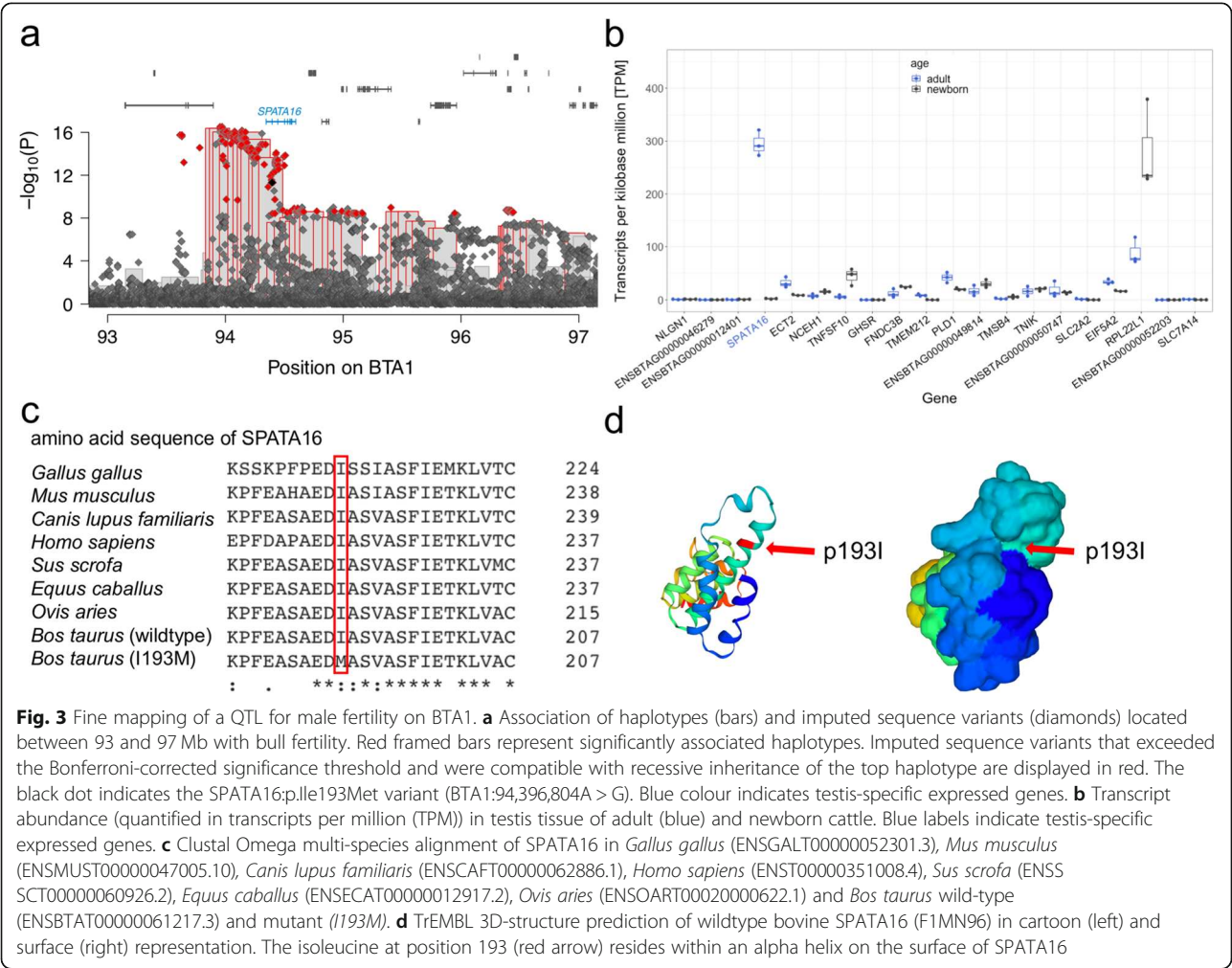
One and 40 sequenced bulls carried the top haplotype in the homozygous and heterozygous state, respectively. Of 44,948 sequence variants that were polymorphic within a window encompassing 3 Mb on either side of

the top haplotype, we detected 764 variants between 93,614,265 and 96,742,540 bp that were compatible with recessive inheritance, i.e. these variants were heterozygous in haplotype carriers and homozygous in the bull that carried the haplotype in the homozygous state. The sequence coverage did not differ between heterozygous, homozygous and non-carrier bulls within that interval.

Of the 764 compatible variants, 505 were also significantly associated with bull fertility at the genome-wide Bonferroni-corrected significance threshold of 3.88×10^{-9} (Fig. 3a). The significantly associated variants clustered between *SPATA16* encoding spermatogenesis associated protein 16 and *NLGN1* encoding neuroligin 1 (Fig. 3a). The most significantly associated variant (BTA1:93,972,058G > A, rs379712951) is in an intron of *NLGN1*, 373 kb upstream the translation start site of *SPATA16*, within the top window from the haplotype-based association study. The *P* value is slightly lower for the BTA1:93,972,058G > A variant than for the most significantly associated haplotype ($P = 3.00 \times 10^{-17}$ vs. 4.16×10^{-17}). We did not consider *NLGN1* as a

candidate gene for an impaired bull fertility, as it is not notably expressed in adult bull testis (TPM < 1). However, we considered *SPATA16* as a positional and functional candidate gene for the BTA1 QTL, because it is testes-specific expressed in cattle (<http://cattlegeneatlas.roslin.ed.ac.uk>) and human (<https://gtexportal.org/home/gene/SPATA16>). Moreover, *SPATA16* mRNA is highly abundant in testis tissue of adult bulls (TPM = 295) (Fig. 3b).

Two of the significantly associated variants compatible with recessive inheritance were in coding regions: a missense variant in *SPATA16* (rs440830663 at BTA1:94,396,804A > G, ENSBTAP00000053460.3: p.Ile193Met, $P = 4.83 \times 10^{-12}$) and a synonymous variant in *GHSR* encoding growth hormone secretagogue receptor (rs714884352, BTA1:95,029,804C > T, ENSBTAP00000014446.5:p.228Leu, $P = 3.7 \times 10^{-9}$). Pathogenic alleles of human *SPATA16* are associated with globozoospermia, i.e., round-headed, often acrosome lacking spermatozoa [25, 26]. The p.Ile193Met variant compatible with recessive inheritance resides



in an evolutionarily conserved tetratricopeptide repeat domain of SPATA16 (Fig. 3c). The isoleucine at position 193 is on the surface of the protein (Fig. 3d). According to protein structure modelling with *Mis-sense3D*, a methionine at position 193 does not alter the tertiary structure of SPATA16. However, it is predicted to be deleterious to SPATA16 function (SIFT score: 0.03, PolyPhen-2 score: 0.998). The *P* value is lower and the effect on male fertility (-0.80 vs. -0.57) is larger for the most significantly associated intergenic than the p.Ile193Met variant in *SPATA16*.

We performed haplotype and sequence-based association studies conditioned on either the top associated haplotype, the SPATA16:p.Ile193Met variant (rs440830663) or intronic BTA1:93,972,058G > A, (rs379712951) to disentangle the QTL (Supplementary Figure 4a,b,c). Conditioning on the top associated haplotype removed the signal in the haplotype study, but not in the sequence-based study (Supplementary Figure 4a). When the SPATA16:p.Ile193Met variant was fixed as covariate in the haplotype and sequence-based association model (Supplementary Figure 4b), the original top haplotype and the variants upstream *SPATA16* were still associated with bull fertility, albeit not at the Bonferroni-corrected significance threshold. The *P* values of the most significantly associated variant (BTA1:93,972,058 bp) and of the top haplotype from the conditional analysis were 1.52×10^{-6} and 1.74×10^{-6} . When the association analysis was conditioned on the BTA1:93,972,058 bp variant, i.e., the most significantly associated variant from the sequence-based association study, the QTL signal was absent in both association studies (Supplementary Figure 4c). In our set of partially imputed sequence variants, BTA1:94,396,804A > G and BTA1:93,972,058G > A were in high linkage disequilibrium ($r^2 = 0.858$).

Coding variants in *ENSBTAG00000006717* and *VWA3A* segregate with the BTA25 QTL

A haplotype-based association analysis conditional on the BTA6 QTL yielded eight significantly associated haplotypes on BTA25 located between 18,975,561 and 20,002,993 bp. The top haplotype (19,628,841 – 19,834,098 bp, $P = 4.56 \times 10^{-8}$) has a frequency of 9% in the BSW population. Despite its strong effect (-1.09 ± 0.20 standard deviations) on bull fertility, the top haplotype was neither associated with routinely examined semen quality nor with sperm morphology features that were assessed as part of the andrological examination.

Of 125 sequenced BSW cattle, 1 and 20 carry the top haplotype in the homozygous and heterozygous state, respectively. Sequence coverage does not differ between homozygous, heterozygous and non-carrier bulls at the BTA25 QTL. We detected 48,121 sequence variants within ± 3 Mb of the top haplotype, of which 778 were compatible with the recessive inheritance of the top

haplotype. Of the compatible variants, six reside in protein-coding sequences: four synonymous variants in *ZP2* (rs110876106, BTA25:19,021,021C > T, p.Pro630%3D), *VWA3A* (rs209406752, BTA25:19,788,285C > T, p.Leu33%3D; rs210416159, BTA25:19,798,888C > T, p.Asp32%3D), and *EEF2K* (rs110751593, BTA25:19,941,219C > T, p.Cys538%3D), a frameshift variant in *ENSBTAG00000006717* (BTA25:19,365,282CA > C, ENSBTAP000000054375.1:p.Ile1167LeufsTer) and a stop gained variant in *VWA3A* (rs434854120, BTA25:19,814,925C > T, ENSBTAP000000021610.1:p.Arg505Ter). Due to a putatively high impact on protein function, we considered the stop gained and frameshift variants in *VWA3A* and *ENSBTAG00000006717* as candidate causal variants for the BTA25 QTL.

ENSBTAG00000006717 encoding the nondescript ATP-binding cassette sub-family A member 3-like protein is highly expressed in testis tissue of adult bulls (127 TPM, Fig. 4b) and shows a testis-specific expression in cattle [27] (<http://cattlegeneatlas.roslin.ed.ac.uk>). The frameshift partly truncates an evolutionarily conserved domain “ATPase associated with a variety of cellular activities (AAA)”. *VWA3A* encoding von Willebrand factor A domain-containing protein 3A is expressed in testis tissue of adult bulls at 17 TPM. The stop mutation truncates *VWA3A* by 58%. The von Willebrand factor A domain-containing protein 3A is a motile cilia-associated protein and might contribute to the beating movement of the flagella [28].

An association study between imputed sequence variants and bull fertility revealed no variants exceeding the Bonferroni-corrected significance threshold (3.88×10^{-9}) at the BTA25 QTL (Fig. 4a). However, 38 variants that did not meet the significance threshold by a small margin ($P < 7.66 \times 10^{-7}$) were between 18,963,390 and 20,033,799 bp. The most significantly associated variant ($P = 4.56 \times 10^{-8}$) was at 19,675,705 bp in an intron of *PDZD9* encoding PDZ domain containing 9. The *P* values were considerably larger for the frameshift and nonsense variant in *ENSBTAG00000006717* ($P = 7.98 \times 10^{-3}$) and *VWA3A* ($P = 5.11 \times 10^{-5}$), respectively.

A missense variant in *ENSBTAG00000019919* segregates with the BTA26 QTL

The top haplotype ($P = 4.9 \times 10^{-17}$) at the BTA26 QTL is between 50,746,717 and 50,993,657 bp. The frequency of the top haplotype in the BSW population is 0.26. Our mapping cohort contained 263 homozygous and 1427 heterozygous haplotype carriers. Of 125 sequenced BSW bulls, 11 and 63 carried the top haplotype in the homozygous and heterozygous state, respectively. The top haplotype was not associated with any other semen quality or sperm morphology trait.

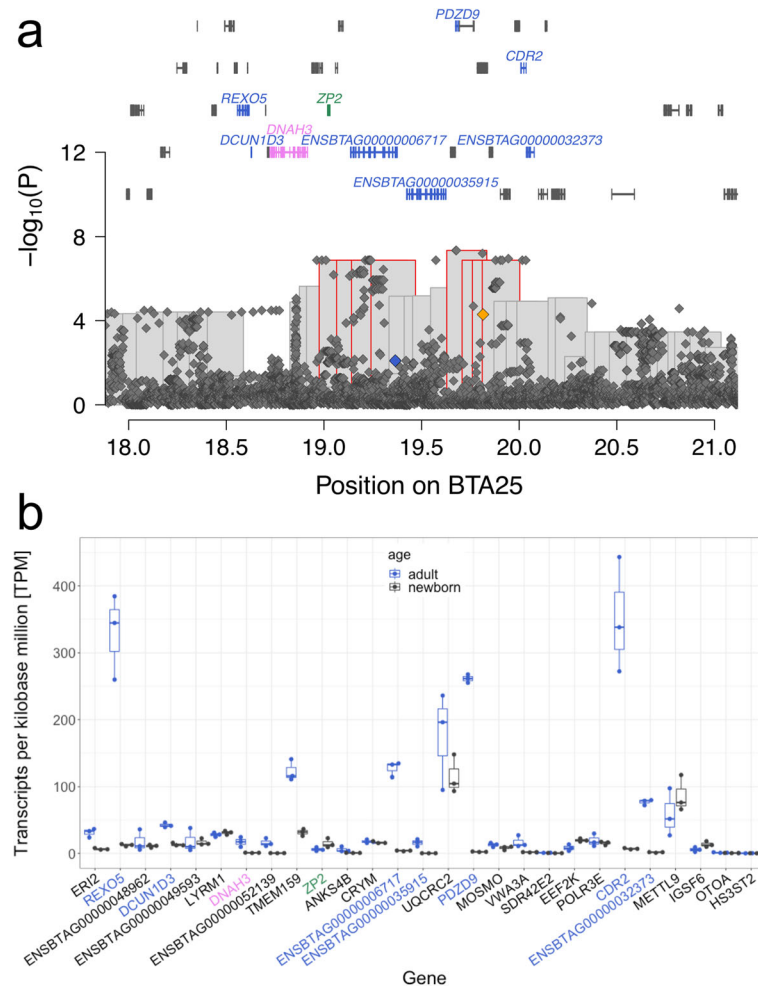


Fig. 4 Detailed view of a QTL for bull fertility on BTA25. **a** Manhattan plot representing the association of imputed sequence variants (diamonds) and haplotypes (bars) with bull fertility. Red framed bars represent haplotypes that met the Bonferroni-corrected significance threshold. The orange and blue diamonds represent the nonsense variant (p.Arg505Ter) in *VWA3A* and frameshift variant (p.Ile1167LeufsTer) in *ENSBTAG00000006717*. Symbols over the Manhattan plot indicate genes and transcripts that are annotated at the QTL region. Genes and transcripts with testis-biased expression, male and female reproductive tract specific expression, and testis and brain specific expression, are indicated in blue, violet and green colour, respectively. **b** Transcript abundance quantified in TPM in testis tissue of adult (blue) and newborn cattle. Genes and transcripts with testis-biased expression, male and female reproductive tract specific expression, and testis and brain specific expression, respectively, are indicated in blue, violet and green colour, respectively

Twenty-six non-coding and two coding variants that were located between 50,748,638 and 50,979,057 bp were compatible with the top haplotype. A synonymous (BTA26: 50,857,552C > T, rs381031945, ENSBTAP00000026536.1: p.831Asp%3D) and a missense variant (BTA26:50,850,915C > G, rs378141069, ENSBTAP00000026536.1: p.Asn576Lys) in *ENSBTAG00000019919* were compatible with the top haplotype. The Ensembl and Refseq annotations indicate that *ENSBTAG00000019919* is the bovine ortholog of *CFAP46* encoding cilia and flagella associated protein 46. The two coding variants in *ENSBTAG00000019919* were in complete linkage disequilibrium and highly significantly associated with bull fertility ($P = 1.3 \times 10^{-15}$) in the sequence-based

association study (Fig. 5a). The p.Asn576Lys-variant is predicted to be deleterious (SIFT: 0.02, PolyPhen-2: 0.921) to protein function and the similarity to the amino acid sequence of other mammalian species is weak in Clustal Omega protein alignment (Fig. 5c). *ENSBTAG00000019919* is expressed at 21 TPM in adult bull testis (Fig. 5b). In *Chlamydomonas reinhardtii*, a CFAP46 homolog is essential for normal flagellar motility [29].

The association study with imputed sequence variant genotypes revealed 18 variants that were stronger associated with bull fertility than the p.Asn576Lys-variant. The two variants with the lowest P value ($P = 1.5 \times 10^{-24}$) were BTA26:50145932C > T (rs720936782) and BTA26:

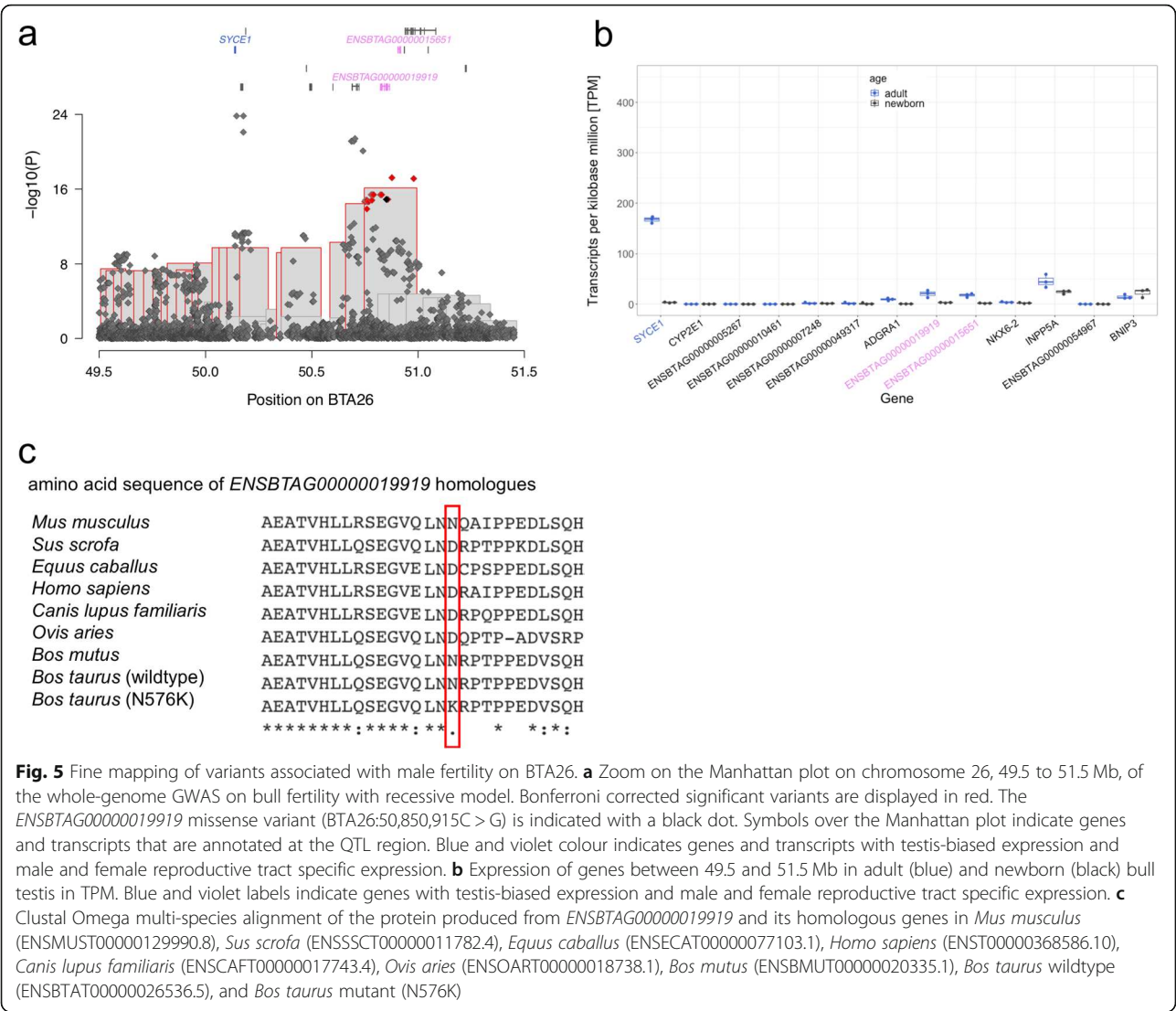


Fig. 5 Fine mapping of variants associated with male fertility on BTA26. **a** Zoom on the Manhattan plot on chromosome 26, 49.5 to 51.5 Mb, of the whole-genome GWAS on bull fertility with recessive model. Bonferroni corrected significant variants are displayed in red. The *ENSBTAG00000019919* missense variant (BTA26:50,850,915C > G) is indicated with a black dot. Symbols over the Manhattan plot indicate genes and transcripts that are annotated at the QTL region. Blue and violet colour indicates genes and transcripts with testis-biased expression and male and female reproductive tract specific expression. **b** Expression of genes between 49.5 and 51.5 Mb in adult (blue) and newborn (black) bull testis in TPM. Blue and violet labels indicate genes with testis-biased expression and male and female reproductive tract specific expression. **c** Clustal Omega multi-species alignment of the protein produced from *ENSBTAG00000019919* and its homologous genes in *Mus musculus* (ENSMUST000000129990.8), *Sus scrofa* (ENSSSCT00000011782.4), *Equus caballus* (ENSECAT00000077103.1), *Homo sapiens* (ENST000000368586.10), *Canis lupus familiaris* (ENSCAFT00000017743.4), *Ovis aries* (ENSOART00000018738.1), *Bos mutus* (ENSBMUT00000020335.1), *Bos taurus* wildtype (ENSBTAT00000026536.5), and *Bos taurus* mutant (N576K)

50178052C > T (rs210673393), that are 4375 bp downstream of the stop codon of *SYCE1* encoding synaptonemal complex central element protein 1 and 2171 bp upstream of the start codon of *CYP2E1* encoding cytochrome P450, family 2, subfamily E, polypeptide (Fig. 5). *SYCE1* is expressed at 167 TPM in adult bull testis, whereas *CYPE1* is not expressed (TPM < 1) (Fig. 5b). However, both variants were not compatible with the inheritance of the top haplotype.

In order to disentangle the QTL, we performed haplotype and sequence-based association studies conditioned on either the top associated haplotype, the p.Asn576Lys-variant in *ENSBTAG00000019919* (BTA26:50,850,915C > G, rs378141069), or the BTA26:50145932C > T (rs720936782) variant downstream *SYCE1* (Supplementary Figure 5a,b,c). The conditional analyses indicated that the top haplotype and the p.Asn576Lys-variant capture the QTL variance only partially, because the

rs720936782 variant was still associated with bull fertility ($P = 1.1 \times 10^{-9}$ and 8.5×10^{-9}), albeit when conditioned on rs378141069 not at the stringent Bonferroni-corrected significance threshold (3.88×10^{-9}). However, the QTL was absent when the analyses were conditioned on rs720936782 (Supplementary Figure 5c). In our set of the imputed sequence variant genotypes, rs378141069 and rs720936782 were in moderate linkage disequilibrium ($r^2 = 0.598$).

Variants compatible with the BTA18 QTL

The most significantly associated haplotype ($P = 5.4 \times 10^{-9}$) for bull fertility on BTA18 was between 36,256,440 and 36,555,965 bp. The haplotype was not associated with any semen quality or sperm morphology trait. The haplotype segregates in the BSW population at a frequency of 0.32. Of 3736 bulls from the mapping cohort,

434 were homozygous and 1690 heterozygous carriers of the top haplotype.

Of 125 BSW bulls with whole-genome sequence variant genotypes, 14 were homozygous and 52 heterozygous haplotype carriers. Sequence variant genotyping revealed 30,018 polymorphic sites between 33,256,440 and 39,555,965 bp, of which 212 were compatible with the recessive inheritance of the top haplotype. Four compatible variants were located in protein-coding regions: a missense variant in *NFAT5* (BTA18:36,768,945, $P = 1.5 \times 10^{-4}$, SIFT score: 0.3 (tolerated)), and three synonymous variants in *ENSBTAG00000052086* (BTA18:36,734,722G > A, rs378665712, $P = 1.5 \times 10^{-4}$), *NQO1* (18:36789835C > T, rs110531779, $P = 1.5 \times 10^{-4}$) and *NOB1* (BTA18:36,816,759A > C, rs41874533, $P = 1.5 \times 10^{-4}$). *NFAT5*, *NQO1*, and *NOB1* are expressed at 22, 16, and 6 TPM in adult bull testes.

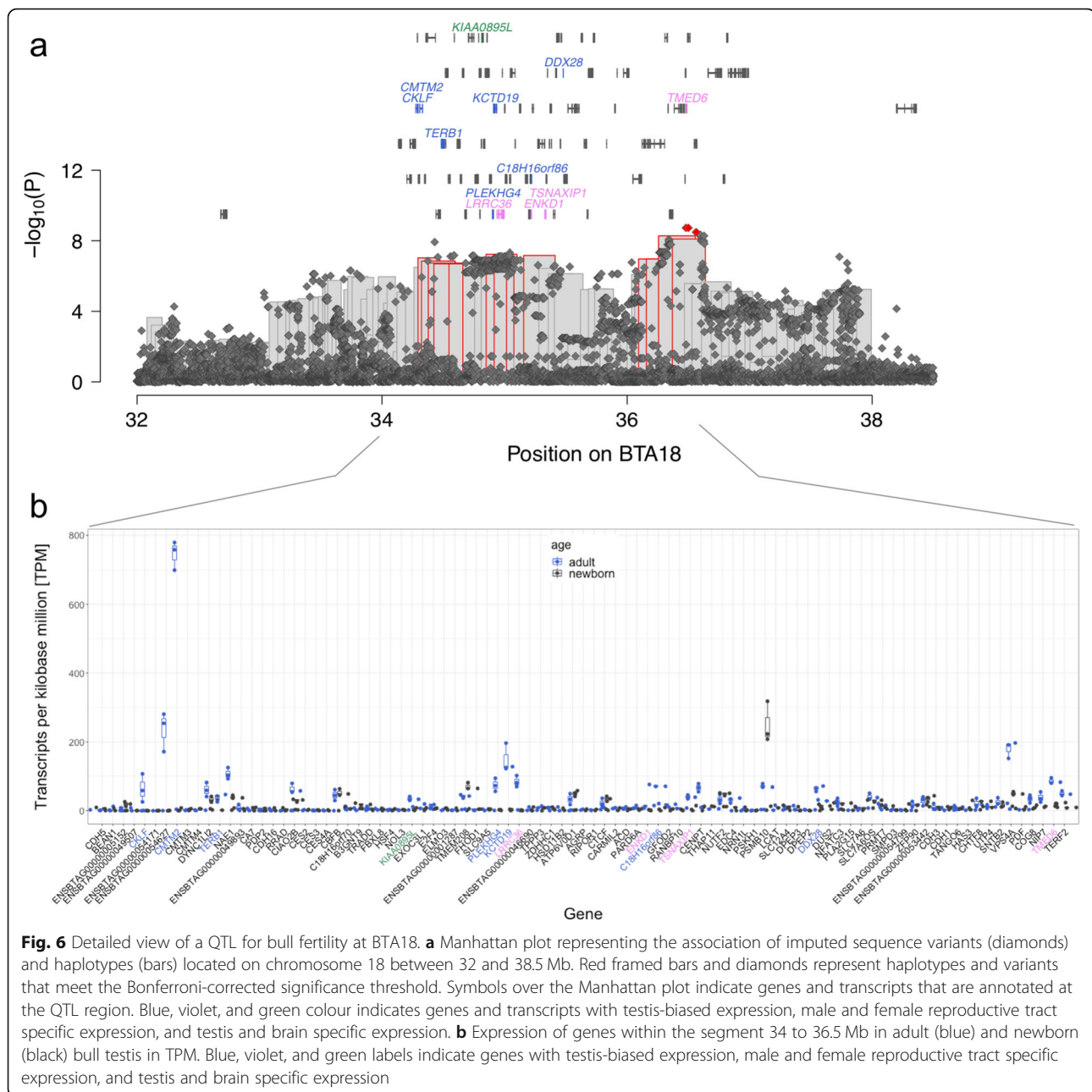
Association testing between imputed sequence variant genotypes and bull fertility revealed four non-coding variants that exceeded the Bonferroni-corrected significance threshold (3.88×10^{-9} , Fig. 6a): a variant considered upstream *NIP7*, downstream *TMEMD6*, and intronic *COG8* (BTA18: 36,480,384 T > G, rs380735496, $P = 1.9 \times 10^{-9}$), an intron variant in *TERF2* (BTA18:36,500,410A > C, rs378216493, $P = 1.9 \times 10^{-9}$), and two intronic variants in *CYB5B* (BTA18:36,561,644 T > A and BTA18:36,565,464A > G, rs136784976 and rs379081158, both $P = 3.3 \times 10^{-9}$). These variants were as well compatible with the recessive inheritance of the top haplotype. *NIP7*, *TMED6*, *COG8*, *TERF2* and *CYB5B* are expressed at 38, 86, 35, 50, and 73 TPM in testis tissue of adult bulls (Fig. 6b).

Discussion

In a mapping cohort of 3736 BSW bulls from which semen was collected at semen collection centers under standardized conditions, only 10% of the phenotypic variation of bull fertility was explained by autosomal variants. This estimate suggests that the additive genetic heritability of bull fertility is low in BSW cattle, which agrees with findings in other cattle populations [30–34]. It is possible that we underestimated the genomic heritability of bull fertility, because the genomic relationship matrix was constructed from autosomal SNPs only. Variants on the sex chromosomes may also contribute to inherited variation in bull fertility [35, 36]. Our mapping cohort provided sufficient statistical power to detect large-effect QTL for a complex trait with an estimated narrow-sense heritability of 10% [37]. However, we did not detect additive QTL. Therefore, we conclude that the additive genetic architecture of bull fertility is polygenic and determined by QTL with small to moderate effects.

Our haplotype-based association analysis identified five recessive QTL with moderate frequencies that, taken together, had a substantial effect on bull fertility in the BSW cattle population. Homozygous haplotype carriers are fertile, but their non-return rate is reduced in artificial inseminations. Our data did not allow investigating bull fertility in natural matings. Phenotypic manifestations of the five recessive QTL detected in our study are less detrimental than effects arising from pathogenic alleles in *CCDC189*, *ARMC3* and *TMEM95* that lead to male sterility in cattle [38–40]. Yet, the hypothetical exclusion of bulls carrying at least one of the five detected top haplotypes in the homozygous state from breeding would significantly increase the 56-day non-return rate in cows and heifers. Thus, the recessive QTL detected in our study contribute to quantitative variation in bull fertility. Similar findings in Fleckvieh, Jersey, and Holstein bulls suggest that the genetic architecture of male reproductive traits is shaped by non-additive effects [16, 20, 40]. Repeatability estimates for traits describing male reproductive performance in cattle are substantially higher than narrow-sense heritability estimates in different cattle breeds including Brown Swiss [22, 41, 42]. In addition to additive effects, the repeatability also includes permanent environmental and non-additive sources of variation. Our study shows that non-additive QTL for bull fertility can be detected in large mapping cohorts using genome-wide association testing. The fact that non-additive effects contribute substantially to the genetic variation of male reproductive traits has implications in efforts to predict bull fertility from dense molecular markers. While improving the fertility of bulls is not a breeding objective in most cattle improvement programs, the genome-based prediction of bull fertility might facilitate implementing compensation strategies for bulls for which the anticipated insemination success is low. Statistical models that consider only additive effects have low predictive power for bull fertility [8, 35], whereas methods that also include non-additive effects may improve the accuracy of genomic predictions [19]. Our findings suggest likewise that non-additive effects should be considered in order to improve the accuracy of genomic prediction for male reproductive traits.

The sequence-based fine-mapping of the QTL revealed candidate causal variants in coding sequences of genes that are either testis-specific expressed in cattle or contribute to variation of male fertility in species other than cattle. The top haplotype on BTA6 is in strong linkage disequilibrium with a likely causal variant that activates cryptic splicing in *WDR19* [22]. A variant in human *WDR19* is associated with morphological abnormalities of sperm and impaired male fertility [43]. The top haplotypes at the BTA1, BTA25 and BTA26 QTL encompass



putatively deleterious variants in genes (*SPATA16*, *VWA3A*, *ENSBTAG00000006717* and *ENSBTAG00000019919*) that are expressed in the bovine male reproductive tract. Furthermore, *SPATA16* and *ENSBTAG00000006717* are specifically expressed in the male reproductive tract (<http://cattlegeneatlas.roslin.ed.ac.uk>) and the produced proteins SPATA16 and ABCA14-like are prevalent in spermatozoa [44, 45]. Pathogenic variants in *SPATA16* and *ENSBTAG00000019919* orthologs have been associated with either impaired male fertility in humans or impaired movement of flagella in a flagellated

green alga [25, 29]. *VWA3A* is associated with motile cilia in humans [28].

Candidate causal variants for recessive traits are frequently prioritized if they are compatible with the segregation of the top haplotype [22, 46–48]. Strictly applying this criterion, the missense and nonsense variants in the coding sequences of *SPATA16*, *VW3A3*, *ENSBTAG00000006717* and *ENSBTAG00000019919* qualify as compelling candidate causal variants for impaired bull fertility; these variants are in linkage disequilibrium with the respective top haplotype and predicted to be deleterious to protein function. Our screen for variants that are compatible with the

inheritance of the top haplotype takes genotyping errors into account. However, due to the low number of sequenced bulls that were homozygous haplotype carriers, it is possible that some variants were misclassified. Thus, we performed a sequence based association study as a complementary fine-mapping approach. An association study with imputed sequence variant genotypes showed that the candidate causal variants are also associated with bull fertility to a certain degree. Yet, non-coding variants were stronger associated with bull fertility than the putatively functional coding variants at all QTL. It is possible that the true causal variants reside in regulatory regions, but are in linkage disequilibrium with the coding variants prioritized in our study. Such a pattern could also indicate multiple trait-associated alleles within a QTL [49].

The p.Asn576Lys variant in *ENSBTAG00000019919* is a particularly compelling candidate causal variant for the BTA26 QTL because it was (i) in linkage disequilibrium with the top haplotype, (ii) highly significantly associated ($P = 1.3 \times 10^{-15}$) in the sequence based association study, (iii) a lysine at position 576 is predicted to be deleterious to protein function, and (iv) the expression of *ENSBTAG00000019919* is restricted to the male and female reproductive tract. However, the sequence-based association study also revealed highly significantly associated variants nearby *SYCE1* that were not in complete linkage disequilibrium with the top haplotype. In fact, an association analysis conditional on the p.Asn576Lys variant did not fully account for the BTA26 QTL variation. The non-coding variant downstream *SYCE1* captured the QTL variance better. *SYCE1* is also a candidate gene for impaired male fertility, because it is testis-specific expressed at very high levels (TPM > 150). Loss of function of *SYCE1* leads to spermatogenic arrest in human and mice [50–52]. The Synaptonemal Complex Central Element Protein 1 is part of the central element of the synaptonemal complex that is required for chromosome synapsis in prophase I of meiosis [53]. Errors in the meiotic recombination can lead to sperm aneuploidy [54] and thus to embryonic loss [55, 56]. It seems plausible that sequence variation affecting *SYCE1* and the p.Asn576Lys variant in *ENSBTAG00000019919* contribute to the BTA26 QTL variation.

It is worth mentioning, that the true causal variants may not necessarily be the most significantly associated variants in association studies with imputed sequence variant genotypes due to imputation errors or sampling effects [57–59]. Interestingly, also the BTA6:58373887C > T variant (rs474302732) that had been reported as a putatively causal variant for the BTA6 QTL [22], was not the top variant in our association study between imputed sequence variant genotypes and bull fertility. Thus, we consider the coding variants in *SPATA16*, *VWA3A*, *ENSBTAG00000006717* and *ENSBTAG00000019919* as

candidate causal variants for impaired male reproductive performance, but further functional investigations and replication studies in independent populations are required to corroborate their association with bull fertility.

Because our association studies considered both bull fertility and semen quality, we were able to reveal the mechanisms through which the BTA1 QTL contributes to variation in fertility. Bulls that are homozygous for the top haplotype encompassing *SPATA16* are less fertile likely because their ejaculates contain an increased number of sperm with morphological abnormalities of the sperm head. Abnormal head morphology is negatively correlated with NRR56 [7, 60]. However, semen quality of homozygous bulls is not as severely compromised as, e.g., in males with globozoospermia [25, 26] and spermatogenic arrest [61], respectively, due to loss-of-function alleles in human and murine *SPATA16*. It is plausible that the *SPATA16*:p.Ile193Met variant compromises semen quality in the homozygous state, as the methionine at position 193 is predicted to be deleterious to *SPATA16* function. However, we can not preclude that non-coding variants in linkage disequilibrium with p.Ile193Met contribute to the BTA1 QTL variation. In any case, the isoleucine residue at position 193 of *SPATA16* is not essential for fertilization, because homozygous bulls are fertile. Sperm head shape abnormalities were only detected in fixed and therefore immotile spermatozoa of homozygous haplotype carriers. Because the morphological abnormalities of the sperm heads are subtle, they remain unnoticed in fresh and motile spermatozoa. In fact, the majority of sperm produced from homozygous haplotype carriers is normal, but the overall proportion of normal sperm is reduced. Many bulls homozygous for the BTA1 QTL are judged to be not suitable for breeding due to insufficient semen quality before their semen is used for artificial inseminations. Thus, the actual effect of the BTA1 QTL on bull fertility is likely more pronounced than observed in our study that was based on insemination success from ejaculates meeting minimum quality requirements.

We found no clues from the analysis of semen quality data why fertility is reduced in bulls that are homozygous carriers of the top haplotypes at the BTA18, BTA25, and BTA26 QTL. Thus, our study confirms indirectly that routine semen quality parameters, once bulls producing ejaculates with clearly insufficient quality are removed, are not sufficient to anticipate bull fertility [62]. Computer-assisted and flow-cytometric sperm analyses may provide further insights into mechanisms that underpin phenotypic variation in bull fertility [11]. Moreover, the phenotype bull fertility does not enable differentiating between fertilization failure and early embryonic losses, because it is based on the 56-day non-return rate. It is thus possible that fertilization is normal,

but the apparently reduced bull fertility is primarily a manifestation of early embryonic losses [63, 64].

Conclusion

Recessive variants contribute substantially to quantitative variation in male fertility in Brown Swiss cattle. Our study suggests that non-additive effects are an important source of variation for traits describing male reproductive performance. Of the five fertility-associated QTL, only two QTL on BTA1 and BTA6 are also associated with reduced semen quality. Mechanisms through which the other QTL impact bull fertility remain unknown. Coding variants are in linkage disequilibrium with the QTL, but it remains difficult to determine their causality because many strongly associated variants in linkage equilibrium reside in poorly annotated non-coding regions.

Methods

Genotypes of BSW bulls

Microarray-derived genotypes of 33,045 BSW cattle (12,887 males and 20,158 females) were provided by breeding associations. The genotypes were determined using various versions of Illumina BovineSNP Bead chips that interrogated between 20 K and 777 K SNPs. Quality control including the removal of variants and samples with more than 10% missing genotypes and the exclusion of variants deviating significantly ($P < 10^{-8}$) from Hardy-Weinberg proportions was carried out using *plink* (version 1.9) [65].

Using a reference panel of 1166 cattle (a subset of the 33,045 BSW) that had Illumina BovineHD BeadChip genotypes for 777,962 SNPs, the genotypes for animals typed at lower density were imputed to higher density (HD) and phased using *Beagle 5.1* [66] with default settings. The final dataset comprised 33,045 animals and 683,609 SNPs. For association testing, we considered only 3881 bulls for which either bull fertility or semen quality data were available.

Bull fertility in Swiss BSW sires

Bull fertility (as at February 2020) was provided by Swissgenetics for 1382 BSW bulls that were used for between 232 and 15,923 first inseminations with conventional (i.e., semen was not sorted for sex) frozen-thawed semen. Bull fertility was estimated based on the 56-day non-return rate after the first insemination in cows and heifers using a linear mixed model similar to the one proposed by Schaeffer et al. [2]: $y_{ijklmnopq} = \mu + MO_i + PA_j + PR_k + (BS \times BK)_{lm} + TE_n + h_o + s_{lp} + e_{ijklmnopq}$, where $y_{ijklmnopq}$ is either 0 or 1, indicating whether or not a subsequent insemination was recorded within a 56-day interval after the first insemination, μ is the intercept, MO_i is the insemination month, PA_j is the parity (heifer

or cow), PR_k is the cost of the semen dose, $(BS \times BK)_{lm}$ is a combination of bull's breed \times cow's breed, TE_n is the insemination technician, h_o is the random effect of the herd, s_{lp} is the random effect of the bull expressed as deviation from the average non-return rate, and $e_{ijklmnopq}$ is a random residual term. Bull fertility was subsequently standardized to a mean of 100 ± 12 . The resulting value is an objective measurement of bull fertility. Three phenotypic outliers (i.e., bulls for which the fertility was more than 5 standard deviations below average) were not considered for subsequent analyses because such bulls might carry rare genetic conditions [40] that can lead to spurious associations [67]. For the GWAS, we considered the 1130 out of the 1382 Swiss BSW bulls which had HD level genotypes that were partially imputed from low-density micro-array derived genotypes. Their fertility was quantified based on 850,708 first artificial inseminations (minimum 232, median 366.5, mean 752.8, and maximum 15,923 per bull).

Bull fertility in German and Austrian BSW sires

Phenotypes for bull fertility were provided by ZuchtData EDV-Dienstleistungen GmbH, Austria, for 4617 BSW bulls from Germany and Austria (as at December 2017) that were used for 4,267,990 and 1,646,254 inseminations in cows and heifers, respectively. Bull fertility was obtained from routine breeding value estimation for reproductive traits, which are jointly estimated for males and females [68]. The statistical model used to estimate breeding values for female reproductive traits included a fixed effect for the service sire that represents bull fertility as deviation from the population mean. We excluded four bulls with exceptionally poor fertility (see above). For association testing, we considered 2606 bulls from the German and Austrian BSW populations which had HD level genotypes that were partially imputed from low-density micro-array derived genotypes. Those bulls were used in a total of 3,746,202 first inseminations (minimum 100, median 360.5, mean 1437.5, maximum 73,858 per bull).

Compilation of a mapping cohort for association testing of bull fertility

To estimate the genetic correlation of bull fertility between the Swiss and German / Austrian bulls, a bivariate restricted maximum likelihood (REML) analysis was performed using the *GCTA* software (version 1.93.2) [69]. Using the "reml-bivar" option in *GCTA* [70], we fitted a genomic relationship matrix that was built using 589,791 autosomal SNPs with minor allele frequency greater than 0.5%. To control for effects attributable to population stratification, we additionally fitted the top 10 principal components of the genomic relationship matrix using the "-qcovar" option [71]. The REML estimation

revealed that the same genetic determinants control the trait in both mapping cohorts ($r_g = 1 \pm 0.28$). We standardized the phenotypes within the Swiss and German / Austrian cohorts to compile a joint mapping population of 3736 bulls. The proportion of phenotypic variation explained by 589,791 autosomal SNPs with minor allele frequency greater than 0.5% was estimated using *GCTA*.

Semen quality of Swiss BSW bulls from routine fresh semen assessment

Semen quality data were provided by Swissgenetics for 74,945 ejaculates that were collected from 1432 BSW bulls between January 2000 and November 2019. Parameters assessed for fresh ejaculates were semen volume (in ml), sperm concentration (million sperm per ml) quantified using photometric analysis, and sperm motility (percentage of sperm with forward motility) assessed visually using a heated-stage microscope at 200-fold magnification. Moreover, the amount of sperm with head and tail anomalies was classified for each ejaculate with scores ranging from 0 to 3 (0: no or very few anomalies, 1: less than 10% sperm with anomalies, 2: between 10 and 30% sperm with anomalies, 3: more than 30% sperm with anomalies). Ejaculates that fulfilled the requirements for artificial insemination (semen volume above 1 ml, more than 300 million sperm per ml, at least 70% motile sperm, no apparent impurities and no excessive amount of sperm with head and tail abnormalities) were diluted using a Tris-egg yolk based extender, filled in straws containing between 15 and 25 million sperm, and cryoconserved in liquid nitrogen. The raw ejaculate data were filtered according to the parameters listed in Table 2. Response variables for the association study

were the mean semen quality values of 902 bulls assessed in at least eight ejaculates per bull.

Sperm morphology data of Swiss BSW bulls from breeding soundness examinations

Sperm morphology was provided by Swissgenetics for 1766 ejaculates of 669 BSW bulls that entered the semen collection center between August 2009 and February 2020. Technicians systematically evaluate sperm morphology of prospective breeding bulls as part of the andrological examination in formalin-fixed fresh semen samples under oil immersion light microscope. Between 150 and 200 sperm per ejaculate are classified according to three non-compensatory and 13 compensatory sperm morphology categories that are considered as either minor or major defects (Table 3). For ejaculates that evidently fail the quality assessment, only 50 spermatozoa are classified. For sperm with multiple morphological abnormalities, only the most severe defect is recorded. Bulls producing ejaculates that contain at least 75% normal sperm and not more than 20% sperm with non-compensatory defects are considered suitable for breeding. If these requirements are not fulfilled, the sperm morphology examination is repeated approximately 4 weeks later. If a bull produces ejaculates that contain between 65 and 74% normal sperm and not more than 20% sperm with non-compensatory defects, its ejaculates are processed but the reduced semen quality is compensated with more sperm per insemination dose. Bulls that repeatedly produce ejaculates with less than 65% normal sperm and more than 20% sperm with non-compensatory defects are not used for breeding.

For the association study, we considered three trait categories of the sperm morphology examination:

Table 2 Parameters and values applied to filter the semen quality data of Swiss BSW bulls

Filtering parameter	Number of bulls	Number of ejaculates
Raw data	1432	74,945
Age at ejaculate collection > 400 and < 1000 days	1380	60,246
Interval to preceding ejaculate known (without first ejaculate)	1379	59,353
Ejaculate volume recorded	1366	58,846
No mixed ejaculates	1358	42,870
Only processed ejaculates	1311	38,370
Interval between ejaculates ≤ 8 days	1304	38,023
Only first ejaculate per day	1301	34,665
Sperm motility ≥ 70%	1285	34,510
Anomaly score plausible	1285	34,505
Sperm per straw ≥ 15 million (not sorted for sex)	1285	34,278
Semen collector and bull handler known	1253	32,023
(imputed) HD genotype available	1119	29,814
≥ 8 records per bull	902	28,856

Table 3 Sperm morphology traits examined in Swiss BSW bulls. Description of sperm morphology traits and their classification into compensatory or non-compensatory and major or minor defects as used in the breeding soundness examination of bulls at the Swiss artificial insemination centre Swissgenetics

Defect number	Description	Compensatory or non-compensatory	Major or minor
0	Abnormal head shape (pyriform / pear shape, tapered / narrow head base, others)	Non-compensatory	Major
1	Vacuoles	Non-compensatory	Major
2	Condensed DNA	Non-compensatory	Major
3	Micro- or macroencephaly	Compensatory	Minor
4	Loose heads	Compensatory	Minor
5	Knobbed acrosome	Compensatory	Major
6	Detached acrosome	Compensatory	Minor
7	Dag defect	Compensatory	Major
8	Curled tail (clef-like)	Compensatory	Major
9	Middle-piece knobbed	Compensatory	Major
10	Tail abaxially attached to head	Compensatory	Minor
11	Tail end loop	Compensatory	Minor
12	Cytoplasmic droplet proximal (middle-piece near head)	Compensatory	Major
13	Cytoplasmic droplet distal (middle-piece near tail)	Compensatory	Minor
14	Underdeveloped	Compensatory	Major
15	Double form	Compensatory	Minor

- The total number of examinations per bull ($n = 595$ bulls with 1577 ejaculates). The number of examined ejaculates ranged from 1 to 26 (median 2).
- The average proportion of sperm for 13 defect categories (Table 3) was quantified for 575 bulls based on 1165 examined ejaculates for which at least 150 spermatozoa were assessed. Between 1 and 13 (median = 1) ejaculates per bull were examined.
- Aggregate phenotypes were the average proportion of sperm with normal morphology (assessed sperm - sum of all morphological defects), major defects, minor defects, compensatory defects and non-compensatory defects (Table 3).

Haplotype-based association testing

Haplotype-based association testing was implemented in R using a sliding window-based approach that was previously applied to map complex and binary traits [22, 72]. We shifted a window of 50 contiguous SNPs in steps of 15 SNPs along the autosomal haplotypes. Within each sliding window, all haplotypes with a frequency above 1% were tested for association with bull fertility or semen quality using the linear model $y = \mu + \sum_{j=1}^{10} a_j PC_j + bHT + ck + e$, where y is a vector of phenotypes (see above), μ is the intercept, PC_j are the top 10 principal components of the genomic relationship matrix, a , b and c are effects of the principal components, the haplotype (HT) tested and a covariable k (e.g., the recessive

BTA6 haplotype associated with male fertility), respectively, and e is a vector of residuals that are assumed to be normally distributed. Haplotypes were tested for association assuming either additive or recessive mode of inheritance. Recessive tests were carried out for haplotypes that were observed in the homozygous state in at least 0.5% of individuals. The genomic relationship matrix was built with *plink* (version 1.9) [65] using 589,791 autosomal SNPs with minor allele frequency greater than 0.5%. Its principle components were calculated with *GCTA* (version 1.9.2.1) [69]. The genomic inflation factor lambda was calculated as: $\lambda = \text{median}(qchisq(p, 1)) / qchisq(0.5, 1)$ where p is a vector of P values.

Redefining a previously detected fertility-associated haplotype on BTA6

We previously detected a recessive QTL for semen quality and bull fertility on BTA6 [22] that segregates at high frequency in BSW cattle. In order to localize the QTL in our mapping cohort, we performed an association study between bull fertility and 6962 haplotypes (31,362 partially imputed SNPs) of chromosome 6 in 1130 Swiss BSW bulls using a haplotype-based association mapping approach (see above) assuming recessive inheritance. The association study revealed the strongest association ($P = 2.4 \times 10^{-20}$) with bull fertility for a haplotype located between 57,628,627 and 57,919,676 bp. In association analyses that were conditioned on the BTA6 QTL, we

fixed the bulls' statuses (recessively coded: non-carrier and heterozygous vs. homozygous) for this haplotype as a covariate.

Whole-genome sequencing, sequence variant genotyping and imputation

We considered 368 BSW and Original Braunvieh (OBV) cattle that were sequenced at between 3 and 60-fold (median of 11.6) genome coverage using either Illumina HiSeq 2500, Illumina HiSeq 4000, or Illumina NovaSeq6000 instruments. The raw sequencing reads were filtered using the *fastp* software [73] with default parameter settings and subsequently aligned to the ARS-UCD1.2 assembly of the bovine genome [74] using the mem-algorithm of the *Burrows Wheeler Aligner (BWA)* software package [75]. Single nucleotide and short insertion and deletion polymorphisms were detected and genotyped using the *HaplotypeCaller*, *GenomicsDBImport* and *GenotypeGVCF* modules of *GATK* [76]. Low quality variants were filtered out using *GATK*'s best practice recommendations for hard filtering. Following filtration, 24,614,962 SNPs and 3,981,732 indels were retained for subsequent analyses. *Beagle 4.1* [77] phasing and imputation was applied to improve the primary genotype calls from *GATK* and infer missing genotypes.

We used the *Beagle 5.1* [66] software with default settings to impute the genotypes for 28,596,694 sequence variants for 3736 bulls from the mapping cohort that had partially imputed genotypes at 683,609 SNPs from a reference panel of 368 sequenced BSW and OBV animals.

Sequence-based association study

Imputed sequence variant genotypes that had allele frequency greater than 0.05 and imputation accuracy (*Beagle* r^2) greater than 0.4 were considered for association testing. The statistical model applied to test for an association between the imputed sequence variants and bull fertility was analogous to the haplotype-based association study. We used the *plink* (version 1.9) [65] software to fit a linear regression model assuming a recessive mode of inheritance conditional on the top ten principal components of a genomic relationship matrix (see above) and the fertility-associated haplotype on BTA6 (see above).

Prioritization of candidate causal variants

Variants that exceeded the Bonferroni-corrected significance threshold of 3.88×10^{-9} (defined based on 12,861,528 imputed sequence variants with $MAF > 0.05$ and *Beagle* $r^2 > 0.40$) and were compatible with the inheritance of the top haplotype were considered as candidate causal variants. To identify variants compatible with the inheritance of the top haplotype, we considered sequence variant genotypes from 125 BSW cattle with bull

fertility phenotypes that had both array-derived genotypes and sequence variant genotypes. Their carrier status for the top haplotype at the QTL was inferred from the array-derived genotypes. We considered sequence variants within ± 3 Mb of the top haplotype as positional candidate causal variants. We filtered for variants that had allele frequency

- greater than 0.8 in animals that carried the top haplotype in the homozygous state,
- between 0.4 and 0.6 in heterozygous haplotype carriers,
- and less than 0.05 in non-carrier animals.

Functional consequences of candidate causal variants were predicted based on the Ensembl (release 101, August 2020) annotation of the ARS-UCD1.2 *Bos taurus* assembly using the *Variant Effect Predictor* (VEP) software tool [78].

Bioinformatic analyses

The *Clustal Omega* [79] software was used for multi-species alignment of protein sequences and *Missense3D* [80] to predict putative effects of missense variants on protein tertiary structure. The three dimensional protein structure was modelled using the *SWISS-MODEL* platform [81].

Transcriptome analysis

To assess the abundance of transcripts in bovine testis, we downloaded between 47 and 58 million 2×150 bp paired-end sequencing reads from the European Nucleotide Archive (ENA, <https://www.ebi.ac.uk/ena>) that were generated using RNA extracted from testis tissue samples of three mature bulls and three newborn male calves of the Angus beef cattle breed (ENA accession numbers: SAMN09205186-SAMN09205191 [82]). The RNA sequencing reads were pseudo-aligned to an index of the bovine transcriptome (ftp://ftp.ensembl.org/pub/release-98/fasta/bos_taurus/cdna/Bos_taurus.ARS-UCD1.2.cdna.all.fa.gz) and transcript abundance was quantified using the *kallisto* software [83]. We used the R package *tximport* [84] to aggregate transcript abundances to the gene level.

Imaging of spermatozoa from bulls homozygous for the haplotype on BTA1

Fresh semen samples were provided by Swissgenetics for two bulls that were homozygous carriers of the fertility-associated haplotype on BTA1. The haplotype status of the bulls was derived using microarray-derived genotypes. Sperm morphology was assessed in semen samples fixed in buffered formaldehyde saline solution using oil-immersion phase contrast microscopy (Olympus BX50, UplanF1 100 \times /1.30, Olympus, Wallisellen, Switzerland).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07523-3>.

Additional file 1. (PDF 16281 kb)

Additional file 2. (TXT 5.35 kb)

Additional file 3. (TXT 18124 kb)

Additional file 4. (TXT 18022 kb)

Additional file 5. (TXT 17920 kb)

Additional file 6. (TXT 17203 kb)

Additional file 7. (TXT 17203 kb)

Additional file 8. (TXT 17612 kb)

Additional file 9. (TXT 17612 kb)

Additional file 10. (TXT 17715 kb)

Additional file 11. (TXT 18022 kb)

Additional file 12. (TXT 18329 kb)

Additional file 13. (TXT 18534 kb)

Additional file 14. (TXT 18534 kb)

Additional file 15. (TXT 18124 kb)

Additional file 16. (TXT 18227 kb)

Additional file 17. (TXT 18227 kb)

Additional file 18. (TXT 18227 kb)

Additional file 19. (TXT 18329 kb)

Additional file 20. (TXT 18022 kb)

Additional file 21. (TXT 18534 kb)

Additional file 22. (TXT 18124 kb)

Additional file 23. (TXT 18022 kb)

Additional file 24. (TXT 18124 kb)

Additional file 25. (TXT 18329 kb)

Additional file 26. (TXT 18534 kb)

Additional file 27. (TXT 17920 kb)

Additional file 28. (TXT 18534 kb)

Additional file 29. (TXT 17920 kb)

Additional file 30. (TXT 18022 kb)

Additional file 31. (TXT 18022 kb)

Additional file 32. (TXT 17920 kb)

Additional file 33. (CSV 1879 kb)

Additional file 34. (CSV 1018 kb)

Additional file 35. (CSV 1135 kb)

Additional file 36. (CSV 432 kb)

Acknowledgements

We appreciate the explanations of sperm morphological examination at Swissgenetics by Susanne Meese and Sarah Wyck. We thank Sarah Wyck and Tjasa Kampara for providing semen for morphological analysis. We acknowledge Swissgenetics for providing phenotype data for the Swiss BSW bulls. We thank Franz Seefried for the identification of homozygous haplotype carriers in the Swiss BSW population. We thank Braunvieh Schweiz for providing genotype data of Swiss BSW bulls. We acknowledge the Arbeitsgemeinschaft Deutsches Braunvieh, Braunvieh Austria, Tierzuchtforschung Grub, the Chair of Animal Breeding of TU München, the Institute of Animal Breeding from Bayerische Landesanstalt für Landwirtschaft and ZuchtData EDV Dienstleistungen GmbH for providing genotype and phenotype data of the Austrian and German BSW bulls. We thank the Functional Genomics Center Zurich for generating DNA sequencing data.

Authors' contributions

MH, NKK and HP analysed and interpreted the data. UW coordinated the collection of semen quality and sperm morphology data. FSH estimated bull fertility for the Swiss BSW bulls. FJ examined and imaged spermatozoa in phase-contrast light microscopy. MH and HP wrote the paper. All authors reviewed and approved the final version of the manuscript.

Funding

This study was financially supported from Swissgenetics, Zollikofen, Switzerland (<https://swissgenetics.ch/>) Grant-ID: 13453, the Förderverein Biotechnologieforschung e.V. (FBF), Bonn, Germany (<https://www.fbf-forschung.de/>) Grant-ID: 2–70024-18, and the Swiss National Science Foundation (310030_185229).

Availability of data and materials

Whole-genome sequence data from 125 BSW cattle are available at the sequence read archive of the European Nucleotide Archive (<https://www.ebi.ac.uk/ena/>), identifiers are given in Supplementary File 2. The results from 29 haplotype-based association studies as well as results from association studies between imputed sequence variants and bull fertility on Chromosomes 1, 18, 25, and 26 are available as Supplementary Data in Supplementary Files 3 to 36. The SNP array data and the phenotype data are available from the organisations mentioned in the Acknowledgement section, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the owners.

Declarations

Ethics approval

No ethics approval was required for the study. Semen samples of bulls were provided from an artificial insemination center after regular collection as part of their insemination straw production. Semen quality, sperm morphology, and insemination success were recorded at the Swiss artificial insemination center Swissgenetics to monitor bull fertility. Fertility records and genotypes were provided by breeding organizations and had been collected as part of normal breeding procedures. No samples were collected specifically for the present study.

Consent for publication

Not applicable.

Competing interests

HP is a member of the editorial board of BMC Genomics. UW and FSH are employees of Swissgenetics. MH, NKK, and FJ declare no financial or non-financial competing interests.

Author details

¹Animal Genomics, Institute of Agricultural Sciences, ETH Zürich, Eschikon 27, 8315 Lindau, Switzerland. ²Clinic of Reproductive Medicine, Vetsuisse Faculty, University of Zurich, 8057 Zurich, Switzerland. ³Swissgenetics, 3052 Zollikofen, Switzerland.

Received: 5 January 2021 Accepted: 5 March 2021

Published online: 30 March 2021

References

1. Thibier M, Wagner H-G. World statistics for artificial insemination in cattle. *Livest Prod Sci.* 2002;74(2):203–12 [https://doi.org/10.1016/S0301-6226\(01\)00291-3](https://doi.org/10.1016/S0301-6226(01)00291-3).
2. Schaeffer LR. Evaluation of bulls for nonreturn rates within artificial insemination organizations. *J Dairy Sci.* 1993;76(3):837–42 [https://doi.org/10.3168/jds.S0022-0302\(93\)77409-3](https://doi.org/10.3168/jds.S0022-0302(93)77409-3).
3. Farrell PB, Presicce GA, Brockett CC, Foote RH. Quantification of bull sperm characteristics measured by computer-assisted sperm analysis (CASA) and the relationship to fertility. *Theriogenology.* 1998;49(4):871–9 [https://doi.org/10.1016/S0093-691X\(98\)00036-3](https://doi.org/10.1016/S0093-691X(98)00036-3).
4. Morrell JM, Nongbua T, Valeanu S, Lima Verde I, Lundstedt-Enkel K, Edman A, et al. Sperm quality variables as indicators of bull fertility may be breed

- dependent. *Anim Reprod Sci.* 2017;185:42–52 <https://doi.org/10.1016/j.anireprosci.2017.08.001>.
5. Saacke RG. Sperm morphology: its relevance to compensable and uncompensable traits in semen. *Theriogenology.* 2008;70(3):473–8 <https://doi.org/10.1016/j.theriogenology.2008.04.012>.
 6. Christensen P, Boelling D, Pedersen KM, Korsgaard IR, Jensen J. Relationship between sperm viability as determined by flow Cytometry and nonreturn rate of dairy bulls. *J Androl.* 2005;26(1):98–106.
 7. Al-Makhzoomi A, Lundeheim N, Håård M, Rodríguez-Martínez H. Sperm morphology and fertility of progeny-tested AI dairy bulls in Sweden. *Theriogenology.* 2008;70(4):682–91 <https://doi.org/10.1016/j.theriogenology.2008.04.049>.
 8. Kastelic JP, Thundathil JC. Breeding soundness evaluation and semen analysis for predicting bull fertility. *Reprod Domest Anim.* 2008;43:368–73 <https://doi.org/10.1111/j.1439-0531.2008.01186.x>.
 9. Vincent P, Underwood SL, Dolbec C, Bouchard N, Kroetsch T, Blondin P. Bovine semen quality control in artificial insemination centers. In: Hopper RM, editor. *Bovine reproduction*. Hoboken: Wiley; 2014. p. 685–95. <https://doi.org/10.1002/9781118833971.ch74>.
 10. Gliozzi TM, Turri F, Manes S, Cassinelli C, Pizzi F. The combination of kinetic and flow cytometric semen parameters as a tool to predict fertility in cryopreserved bull semen. *Animal.* 2017;11:1975–82.
 11. Sellem E, Broekhuijsen MLWJ, Chevrier L, Camugli S, Schmitt E, Schibler L, et al. Use of combinations of in vitro quality assessments to predict fertility of bovine semen. *Theriogenology.* 2015;84:1447–1454.e5.
 12. Utt MD. Prediction of bull fertility. *Anim Reprod Sci.* 2016;169:37–44 <https://doi.org/10.1016/j.anireprosci.2015.12.011>.
 13. Bucher K, Malama E, Siuda M, Janett F, Bollwein H. Multicolor flow cytometric analysis of cryopreserved bovine sperm: a tool for the evaluation of bull fertility. *J Dairy Sci.* 2019;102(12):11652–69 <https://doi.org/10.3168/jds.2019-16572>.
 14. Ibanescu I, Siuda M, Bollwein H. Motile sperm subpopulations in bull semen using different clustering approaches – associations with flow cytometric sperm characteristics and fertility. *Anim Reprod Sci.* 2020;215:106329 <https://doi.org/10.1016/j.anireprosci.2020.106329>.
 15. Druet T, Fritz S, Sellem E, Basso B, Gérard O, Salas-Cortes L, et al. Estimation of genetic parameters and genome scan for 15 semen characteristics traits of Holstein bulls. *J Anim Breed Genet.* 2009;126(4):269–77 <https://doi.org/10.1111/j.1439-0388.2008.00788.x>.
 16. Nani JP, Peñagaricano F. Whole-genome homozygosity mapping reveals candidate regions affecting bull fertility in US Holstein cattle. *BMC Genomics.* 2020;21(1):338 <https://doi.org/10.1186/s12864-020-6758-y>.
 17. Nicolini P, Amorín R, Han Y, Peñagaricano F. Whole-genome scan reveals significant non-additive effects for sire conception rate in Holstein cattle. *BMC Genet.* 2018;19(1):14 <https://doi.org/10.1186/s12863-018-0600-4>.
 18. Peñagaricano F, Weigel KA, Khatib H. Genome-wide association study identifies candidate markers for bull fertility in Holstein dairy cattle. *Anim Genet.* 2012;43:65–71 <https://doi.org/10.1111/j.1365-2052.2012.02350.x>.
 19. Nani JP, Rezende FM, Peñagaricano F. Predicting male fertility in dairy cattle using markers with large effect and functional annotation data. *BMC Genomics.* 2019;20(1):258 <https://doi.org/10.1186/s12864-019-5644-y>.
 20. Rezende FM, Dietsch GO, Peñagaricano F. Genetic dissection of bull fertility in US Jersey dairy cattle. *Anim Genet.* 2018;49(5):393–402 <https://doi.org/10.1111/age.12710>.
 21. Ferencaković M, Šolknjer J, Kapš M, Curik I. Genome-wide mapping and estimation of inbreeding depression of semen quality traits in a cattle population. *J Dairy Sci.* 2017;100(6):4721–30 <https://doi.org/10.3168/jds.2016-12164>.
 22. Hiltpold M, Niu G, Kadri NK, Crysnanto D, Fang Z-H, Spengeler M, et al. Activation of cryptic splicing in bovine WDR19 is associated with reduced semen quality and male fertility. *PLoS Genet.* 2020;16(5):e1008804 <https://doi.org/10.1371/journal.pgen.1008804>.
 23. Khatib H, Huang W, Wang X, Tran AH, Bindrim AB, Schutzkus V, et al. Single gene and gene interaction effects on fertilization and embryonic survival rates in cattle. *J Dairy Sci.* 2009;92(5):2238–47 <https://doi.org/10.3168/jds.2008-1767>.
 24. Khayatizadeh N, Mészáros G, Utsunomiya YT, Schmitz-Hsu F, Seefried F, Schnyder U, et al. Genome-wide mapping of the dominance effects based on breed ancestry for semen traits in admixed Swiss Fleckvieh bulls. *J Dairy Sci.* 2019;102(12):12117–24 <https://doi.org/10.3168/jds.2019-16899>.
 25. Dam AHDM, Kosciński I, Kremer JAM, Moutou C, Jaeger A-S, Oudakker AR, et al. Homozygous mutation in SPATA16 is associated with male infertility in human Globozoospermia. *Am J Hum Genet.* 2007;81(4):813–20 <https://doi.org/10.1086/521314>.
 26. Ellnati E, Fossard C, Okutman O, Ghédri H, Ibalá-Romdhane S, Ray PF, et al. A new mutation identified in SPATA16 in two globozoospermic patients. *J Assist Reprod Genet.* 2016;33(6):815–20 <https://doi.org/10.1007/s10815-016-0715-3>.
 27. Fang L, Cai W, Liu S, Canela-Xandri O, Gao Y, Jiang J, et al. Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. *Genome Res.* 2020;30:790–801 <https://doi.org/10.1101/gr.250704.119>.
 28. Patir A, Fraser AM, Barnett MW, McTeir L, Rainger J, Davey MG, et al. The transcriptional signature associated with human motile cilia. *Sci Rep.* 2020;10(1):10814 <https://doi.org/10.1038/s41598-020-66453-4>.
 29. Brown JM, DiPetrillo CG, Smith EF, Witman GB. A FAP46 mutant provides new insights into the function and assembly of the C1d complex of the ciliary central apparatus. *J Cell Sci.* 2012;125(16):3904–13 <https://doi.org/10.1242/jcs.107151>.
 30. Berry DP, Wall E, Pryce JE. Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal.* 2014;8:105–21.
 31. Boichard D, Manfredi E. Genetic analysis of conception rate in French Holstein cattle. *Acta Agriculturae Scandinavica Section A Animal Science.* 1994;44(3):138–45 <https://doi.org/10.1080/090647094090410890>.
 32. Kuhn MT, Hutchison JL. Prediction of dairy bull fertility from field data: use of multiple services and identification and utilization of factors affecting bull fertility. *J Dairy Sci.* 2008;91(6):2481–92 <https://doi.org/10.3168/jds.2007-0743>.
 33. Stålhammar E-M, Janson L, Philipsson J. Genetic studies on fertility in AI bulls. II. Environmental and genetic effects on non-return rates of young bulls. *Anim Reprod Sci.* 1994;34(3-4):193–207 [https://doi.org/10.1016/0378-4320\(94\)90016-7](https://doi.org/10.1016/0378-4320(94)90016-7).
 34. Tiezzi F, Maltecca C, Penasa M, Cecchinato A, Bittante G. Short communication: genetic analysis of dairy bull fertility from field data of Brown Swiss cattle. *J Dairy Sci.* 2013;96(11):7325–8 <https://doi.org/10.3168/jds.2013-6885>.
 35. Fortes MRS, Porto-Neto LR, Satake N, Nguyen LT, Freitas AC, Melo TP, et al. X chromosome variants are associated with male fertility traits in two bovine populations. *Genet Sel Evol.* 2020;52(1):46 <https://doi.org/10.1186/s12711-020-00563-5>.
 36. Pacheco HA, Rezende FM, Peñagaricano F. Gene mapping and genomic prediction of bull fertility using sex chromosome markers. *J Dairy Sci.* 2020;103(4):3304–11 <https://doi.org/10.3168/jds.2019-17767>.
 37. Goddard ME, Hayes BJ. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet.* 2009;10(6):381–91 <https://doi.org/10.1038/nrg2575>.
 38. Iso-Touru T, Wurmser C, Venhoranta H, Hiltpold M, Savolainen T, Sironen A, et al. A splice donor variant in CCDC189 is associated with asthenospermia in Nordic red dairy cattle. *BMC Genomics.* 2019;20(1):286 <https://doi.org/10.1186/s12864-019-5628-y>.
 39. Pausch H, Venhoranta H, Wurmser C, Hakala K, Iso-Touru T, Sironen A, et al. A frameshift mutation in ARMC3 is associated with a tail stump sperm defect in Swedish red (*Bos taurus*) cattle. *BMC Genet.* 2016;17(1):49 <https://doi.org/10.1186/s12863-016-0356-7>.
 40. Pausch H, Kölle S, Wurmser C, Schwarzenbacher H, Emmerling R, Jansen S, et al. A nonsense mutation in TMEM95 encoding a nondescript Transmembrane protein causes idiopathic male subfertility in cattle. *PLoS Genet.* 2014;10(1):e1004044 <https://doi.org/10.1371/journal.pgen.1004044>.
 41. Berry DP, Eivers B, Dunne G, McParland S. Genetics of bull semen characteristics in a multi-breed cattle population. *Theriogenology.* 2019;123:202–8 <https://doi.org/10.1016/j.theriogenology.2018.10.006>.
 42. Olsen HB, Heringstad B, Klemetsdal G. Genetic analysis of semen characteristic traits in young Norwegian red bulls. *J Dairy Sci.* 2020;103(1):545–55 <https://doi.org/10.3168/jds.2019-17291>.
 43. Ni X, Wang J, Lv M, Liu C, Zhong Y, Tian S, et al. A novel homozygous mutation in WDR19 induces disorganization of microtubules in sperm flagella and nonsyndromic asthenoteratospermia. *J Assist Reprod Genet.* 2020;37(6):1431–9 <https://doi.org/10.1007/s10815-020-01770-1>.
 44. Byrne K, Leahy T, McCulloch R, Colgrave ML, Holland MK. Comprehensive mapping of the bull sperm surface proteome. *PROTEOMICS.* 2012;12(23-24):3559–79 <https://doi.org/10.1002/pmic.201200133>.
 45. Lu L, Lin M, Xu M, Zhou Z-M, Sha J-H. Gene functional research using polyethylenimine-mediated in vivo gene transfection into mouse spermatogenic cells. *Asian J Androl.* 2006;8(1):53–9 <https://doi.org/10.1111/j.1745-7262.2006.00089.x>.

46. Fritz S, Capitan A, Djari A, Rodriguez SC, Barbat A, Baur A, et al. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. *PLoS One*. 2013;8(6): e65550 <https://doi.org/10.1371/journal.pone.0065550>.
47. Hoff JL, Decker JE, Schnabel RD, Taylor JF. Candidate lethal haplotypes and causal mutations in Angus cattle. *BMC Genomics*. 2017;18(1):799 <https://doi.org/10.1186/s12864-017-4196-2>.
48. Pausch H, Schwarzenbacher H, Burgstaller J, Flisikowski K, Wurmser C, Jansen S, et al. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. *BMC Genomics*. 2015;16(1):312 <https://doi.org/10.1186/s12864-015-1483-7>.
49. Bickel RD, Kopp A, Nuzhdin SV. Composite effects of polymorphisms near multiple regulatory elements create a major-effect QTL. *PLoS Genet*. 2011;7(1):e1001275 <https://doi.org/10.1371/journal.pgen.1001275>.
50. Bolcun-Filas E, Speed R, Taggart M, Grey C, de Massy B, Benavente R, et al. Mutation of the mouse Syce1 gene disrupts synapsis and suggests a link between Synaptonemal complex structural components and DNA repair. *PLoS Genet*. 2009;5(2):e1000393 <https://doi.org/10.1371/journal.pgen.1000393>.
51. Maor-Sagie E, Cinnamon Y, Yaacov B, Shaag A, Goldsmidt H, Zenvirt S, et al. Deleterious mutation in SYCE1 is associated with non-obstructive azoospermia. *J Assist Reprod Genet*. 2015;32(6):887–91 <https://doi.org/10.1007/s10815-015-0445-y>.
52. Pashaei M, Rahimi Bidgoli MM, Zare-Abdollahi D, Najmabadi H, Haji-Seyed-Javadi R, Fatehi F, et al. The second mutation of SYCE1 gene associated with autosomal recessive nonobstructive azoospermia. *J Assist Reprod Genet*. 2020;37(2):451–8 <https://doi.org/10.1007/s10815-019-01660-1>.
53. Costa Y, Speed R, Öllinger R, Alsheimer M, Semple CA, Gautier P, et al. Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the Centre of meiosis. *J Cell Sci*. 2005;118(12):2755–62 <https://doi.org/10.1242/jcs.02402>.
54. Tempest HG. Meiotic recombination errors, the origin of sperm aneuploidy and clinical recommendations. *Syst Biol Reprod Med*. 2011;57(1-2):93–101 <https://doi.org/10.3109/19396368.2010.504879>.
55. Findikli N, Kahraman S, Saglam Y, Beyazyurek C, Sertyel S, Karlikaya G, et al. Embryo aneuploidy screening for repeated implantation failure and unexplained recurrent miscarriage. *Reprod BioMed Online*. 2006;13(1):38–46 [https://doi.org/10.1016/S1472-6483\(10\)62014-7](https://doi.org/10.1016/S1472-6483(10)62014-7).
56. Turner KJ, Silvestri G, Black DH, Dobson G, Smith C, Handyside AH, et al. Karyomapping for simultaneous genomic evaluation and aneuploidy screening of preimplantation bovine embryos: the first live-born calves. *Theriogenology*. 2019;125:249–58 <https://doi.org/10.1016/j.theriogenology.2018.11.014>.
57. Barendse W. The effect of measurement error of phenotypes on genome wide association studies. *BMC Genomics*. 2011;12(1):232 <https://doi.org/10.1186/1471-2164-12-232>.
58. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brøndum RF, et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat Genet*. 2014;46(8):858–65 <https://doi.org/10.1038/ng.3034>.
59. Pausch H, MacLeod IM, Fries R, Emmerling R, Bowman PJ, Daetwyler HD, et al. Evaluation of the accuracy of imputed sequence variant genotypes and their utility for causal variant detection in cattle. *Genet Sel Evol*. 2017;49(1): 24 <https://doi.org/10.1186/s12711-017-0301-x>.
60. Söderquist L, Janson L, Larsson K, Einarsson S. Sperm morphology and Fertility in A. I. Bulls. *J Vet Med A*. 1991;38(1-10):534–43 <https://doi.org/10.1111/j.1439-0442.1991.tb01045.x>.
61. Fujihara Y, Oji A, Larasati T, Kojima-Kita K, Ikawa M. Human Globozoospermia-related gene Spata16 is required for sperm formation revealed by CRISPR/Cas9-mediated mouse models. *Int J Mol Sci*. 2017;18(10):2208 <https://doi.org/10.3390/ijms18102208>.
62. Gredler B, Fuerst C, Fuerst-Waltl B, Schwarzenbacher H, Sölkner J. Genetic parameters for semen production traits in Austrian dual-purpose Simmental bulls. *Reprod Domest Anim*. 2007;42(3):326–8 <https://doi.org/10.1111/j.1439-0531.2006.00778.x>.
63. Charlier C, Li W, Harland C, Littlejohn M, Coppieters W, Creagh F, et al. NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. *Genome Res*. 2016;26(10):1333–41 <https://doi.org/10.1101/gr.207076.116>.
64. Inskeep EK, Dailey RA. Embryonic death in cattle. *Vet Clin*. 2005;21(2):437–61 <https://doi.org/10.1016/j.cvfa.2005.02.002>.
65. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4(1):7 <https://doi.org/10.1186/s13742-015-0047-8>.
66. Browning BL, Zhou Y, Browning SR. A one-penny imputed genome from next-generation reference panels. *Am J Hum Genet*. 2018;103(3):338–48 <https://doi.org/10.1016/j.ajhg.2018.07.015>.
67. Auer PL, Reiner AP, Leal SM. The effect of phenotypic outliers and non-normality on rare-variant association testing. *Eur J Hum Genet*. 2016;24(8): 1188–94 <https://doi.org/10.1038/ejhg.2015.270>.
68. Fuerst C, Gredler B. Genetic evaluation for fertility traits in Austria and Germany. *Interbull Bull*. 2009;40:3–3 <https://journal.interbull.org/index.php/ib/article/view/1076>.
69. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88(1):76–82 <https://doi.org/10.1016/j.ajhg.2010.11.011>.
70. Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*. 2012;28(19):2540–2 <https://doi.org/10.1093/bioinformatics/bts474>.
71. Vinkhuyzen AAE, Pedersen NL, Yang J, Lee SH, Magnusson PKE, Iacono WG, et al. Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Transl Psychiatry*. 2012;2:e102.
72. Pausch H, Ammermüller S, Wurmser C, Hamann H, Tetens J, Drögemüller C, et al. A nonsense mutation in the COL7A1 gene causes epidermolysis bullosa in Vorderwald cattle. *BMC Genet*. 2016;17(1):149 <https://doi.org/10.1186/s12863-016-0458-2>.
73. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–90 <https://doi.org/10.1093/bioinformatics/bty560>.
74. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience*. 2020;9(3) <https://doi.org/10.1093/gigascience/giaa021>.
75. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-bio]*. 2013. <http://arxiv.org/abs/1303.3997>. Accessed 4 Nov 2019.
76. DePristo MA, Banks E, Poplin RE, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491–8 <https://doi.org/10.1038/ng.806>.
77. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*. 2007;81(5):1084–97 <https://doi.org/10.1086/521987>.
78. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122 <https://doi.org/10.1186/s13059-016-0974-4>.
79. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal omega. *Mol Syst Biol*. 2011;7(1):539 <https://doi.org/10.1038/msb.2011.75>.
80. Ittisoponpisan S, Islam SA, Khanna T, Alhuzimi E, David A, Sternberg MJE. Can predicted protein 3D structures provide reliable insights into whether missense variants are disease associated? *J Mol Biol*. 2019;431(11):2197–212 <https://doi.org/10.1016/j.jmb.2019.04.009>.
81. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46(W1):W296–303 <https://doi.org/10.1093/nar/gky427>.
82. Gao Y, Li S, Lai Z, Zhou Z, Wu F, Huang Y, et al. Analysis of long non-coding RNA and mRNA expression profiling in immature and mature bovine (*Bos taurus*) testes. *Front Genet*. 2019;10 <https://doi.org/10.3389/fgene.2019.00646>.
83. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34(5):525–7 <https://doi.org/10.1038/nbt.3519>.
84. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*. 2016;4 <https://doi.org/10.12688/f1000research.7563.2>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.